TURING 图灵数学·统计学丛书 22

PEARSON Prentice Hall



Operations Research An Introduction

运筹学导论

初级篇

(第8版)

[美] Hamdy A. Taha 著

薛毅 刘德刚 朱建明 侯思祥 译

韩继业 审校



人民邮电出版社
POSTS & TELECOM PRESS

运筹学导论：初级篇（第8版）

Operations Research: An Introduction

"本书全面地论述了运筹学的三个方面——理论、应用和计算，而且游刃有余。我求学时就是通过本书老版本学习运筹学的，如今我在使用新版本教授学生，这么多年了，它仍然是本领域的经典。"

—— Amazon.com

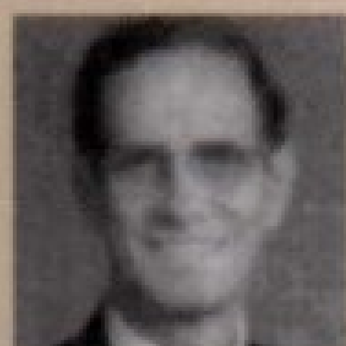
运筹学是一门应用领域十分广泛的学科，它应用分析、试验、量化的方法，对经济管理系统中人力、物力、财力等资源进行统筹安排，为决策者提供有依据的最佳方案，以实现最有效的管理。

本书是运筹学方面的经典著作之一，理论严密，案例丰富，并且充分运用了计算机软件，体现了作者在运筹学教学研究和业界实践方面精湛的造诣，已被翻译成中、韩、西班牙、日、俄、土耳其、印尼、马来等多种语言，为全球众多高校采用，深受好评。第8版对教材内容作了较大的修订，在教材的编排上突出反映运筹学中的应用问题和计算方法。

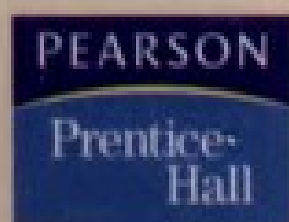
本书特色

- 理论联系实际，应用色彩浓厚。
- 注重与计算机软件程序相结合，富有时代气息。
- 每章开头都有本章导读，帮助读者了解教材内容。

将原书分成两册出版后，对原书章节顺序进行了调整。初级篇内容全面，符合国内的大纲要求，可作为相关专业本科生教材。高级篇适合作为研究生或MBA教学用书和参考读物。



Hamdy A. Taha 美国阿肯色大学荣休教授，世界知名运筹学家，曾在全球各地任教和担任顾问，同时拥有非常丰富的教学研究和实践经验。他在*Management Science*和*Operations Research*等世界顶级学术刊物上发表了大量论文。





www.PearsonEd.com

本书相关信息请访问：图灵网站 http://www.turingbook.com
读者/作者热线：(010) 88593802
反馈/投稿/推荐信箱：contact@turingbook.com

分类建议 数学/应用数学

人民邮电出版社网址 www.ptpress.com.cn



ISBN 978-7-115-18150-3/O1

定价：69.00 元

TURING 图灵数学·统计学丛书22



Operations Research
An Introduction

运筹学导论
初级篇
（第8版）

［美］ Hamdy A. Taha 著

薛毅 刘德刚 朱建明 侯思祥 译

韩继业 审校

人民邮电出版社

北 京

图书在版编目(CIP)数据

运筹学导论：第 8 版，初级篇/(美)塔哈(Taha, H. A.)
著；薛毅等译. —北京：人民邮电出版社，2008. 8
(图灵数学·统计学丛书)
书名原文：Operations Research: An Introduction
ISBN 978-7-115-18150-3

Ⅰ. 运… Ⅱ. ①塔… ②薛… Ⅲ. 运筹学 Ⅳ. O22

中国版本图书馆 CIP 数据核字(2008)第 071809 号

内 容 提 要

本书是运筹学方面的经典著作之一，为全球众多高校采用. 初级篇共 12 章，内容包括线性规划建模、单纯形方法和灵敏度分析、对偶性和后最优分析、运输模型及其变型、网络模型、目标规划、整数线性规划、确定性动态规划、确定性库存模型、决策分析和对策论、排队系统等，并附有 AMPL 建模语言简介.

本书可作为经营类专业、数学专业和计算机专业本科生的教材，也可供相关研究人员参考.

图灵数学·统计学丛书

运筹学导论：初级篇(第 8 版)

◆ 著　[美] Hamdy A. Taha
译　薛　毅　刘德刚　朱建明　侯思祥
审校　韩继业
责任编辑　张继发

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
邮编 100061　电子函件 315@ptpress.com.cn
网址：http://www.ptpress.com.cn
北京隆昌伟业印刷有限公司印刷

◆ 开本：700×1000　1/16
印张：33
字数：726 千字　2008 年 8 月第 1 版
印数：1–3 000 册　2008 年 8 月北京第 1 次印刷

著作权合同登记号　图字：01-2006-5777 号

ISBN 978-7-115-18150-3/O1

定价：69.00 元

读者服务热线：(010)88593802　印装质量热线：(010) 67129223

反盗版热线：(010)67171154

版权声明

译 者 序

运筹学起源于 20 世纪二次大战期间, 是一门应用性很强的学科. 1938 年, 英国皇家空军部门在 Bawdsey 成立了一个从事作战研究的科学家小组, 小组的科学家把他们的研究工作称为 “operational research” (“operation” 在军事术语中意为 “作战”). 这是 “运筹学” 一词最早出现于文献的时间. 二战中英军每一个大的指挥部大都成立了这种运筹研究小组. 之后, 美国和加拿大的军事部门也成立了若干运筹研究小组 (美国称这种研究工作为 “operations research”). 他们广泛地研究有关战果评价、战术革新、技术援助、战略选择和战术计划等问题. 二战期间英、美、加等国军事部门的运筹研究小组的工作为同盟国战胜德、意、日等轴心国做出了卓越的贡献. 但对于人类社会的科学进程而言, 这些科学家的集体工作和智慧开创了一门崭新的学科 —— 运筹学.

体现运筹学思想和方法的某些早期先驱性的研究工作, 可以追溯到 20 世纪初期. 例如, 1908 年丹麦工程师埃尔朗提出的电话话务理论是运筹学中排队论 (queueing theory) 的起源; 1916 年英国的兰彻斯特提出的战斗模型方程是军事运筹学早期的一项重要成果; 1939 年前苏联数学家坎托罗维奇在 *The Mathematical Method of Production Planning and Organization* 一书中, 开创性地提出线性规划, 并研究了工业生产的资源合理利用和计划等问题, 这一卓越贡献使他获得了 1975 年诺贝尔经济学奖; 基本的对策均衡的思想可追溯到 1838 年库尔诺的文章; 1913 年德国的策梅洛提出了抽象战略对策的数学模型; 1928 年冯 · 诺伊曼提出了关于二人零和对策的解的一般理论. 这些都是关于对策论的早期研究. 上述这些先驱性成就对以后运筹学的发展有着深远的影响.

二战以后, 美国等国家的军事部门保留和调整了运筹研究组, 人员编制得到了扩大, 运筹学有了新的发展. 1949 年美国成立了著名的兰德 (RAND) 公司. 与此同时, 许多运筹学工作者从军方转入企业、大学或政府部门. 在新的更广阔的环境中, 运筹学的应用研究和理论研究得到了迅速发展, 多年来它已为欧美等国创造了数以亿计的社会财富.

简略地说, 运筹学的研究对象是现实世界中的运行系统, 这些运行系统的设计和运转受到管理人员决策的影响和作用. 运筹学创造出一些理论 (包括数学模型) 和方法, 用来描述和分析这些运行系统的现象、性质和变化, 以寻求影响和作用于运行系统的设计与运转的最有效 (最优) 的决策, 发挥有限资源的最大效益, 使得运行系统达到总体最优的目标.

半个世纪以来, 运筹学在研究和解决各种复杂的实际问题中不断地得到创新和发展, 新模型、新理论和新方法不断涌现, 至今它已成为一个庞大的学科, 包括线性的和非线性的、连续的和离散的、确定性的和不确定性的许多分支. 运筹学的基本方法中有数学方法、统计学方法、模拟 (仿真) 方法、计算机科学方法等, 其中各种

优化方法处于非常重要的地位.

运筹学的非凡价值使得许多国家的大学里的运筹学系、管理科学系、经济学系、工业工程系、系统科学系、数学系、计算科学系等早已开设了关于运筹学及其一些分支学科的课程. 我国的情况也大致如此. 从适应于不同院系专业的学生学习运筹学的角度来考虑, 一本好的运筹学基础的教科书十分重要. 在国外关于运筹学基础的诸多教材中, Hamdy A. Taha 著的*Operations Research, An Introduction* 是非常优秀的一本. Taha 是美国阿肯色大学工业学院工业工程教授, 世界知名运筹学家. 他著的*Operations Research, An Introduction, Preliminary edition* 出版于 1968 年. 此书经过多次修改与扩增, 今年已出版了第 8 版. 该书被世界上多所大学采用为运筹学基础教材, 已有西班牙文、日文、俄文、土耳其文、印尼文等多种译本出版. 它有 3 大重要特色. 一是内容广泛、取材得当. 连同附带的光盘, 共有 24 章和 5 个附录, 内容涉及线性规划、运输问题、网络问题、目标规划、整数规划、动态规划、库存问题、非线性规划等确定性运筹分支, 以及随机动态规划、随机库存问题、排队系统、马尔可夫决策过程、决策分析、对策论、模拟问题、预报问题等随机性运筹分支. 这些内容覆盖了迄今运筹学所研究的大部分重要问题. 该书在取材上首先是重视对上述运筹问题的基本知识的讲解, 但对某些问题也包括了较高深的内容, 以满足不同读者的需要. 二是突出实用性. 书中各章总是通过若干实际问题的求解来引导出所要讲的运筹问题的数学模型. 这既凸显出这些运筹问题的实际背景, 也使读者可学到如何进行建模. 第 24 章详细地介绍了 15 个实际应用案例, 运用了多种运筹学技术进行建模、数据采集以及求解计算等. 附录 E 还收录了近 50 个应用例子. 作者精心收集和分析的这些实例来源于工业、商业、金融、社会、体育、娱乐等许多行业, 是很好的运筹学教学资料. 三是计算方法与软件相结合. 全书大量使用教学辅助软件 TORA、电子表格程序 Excel 及 AMPL 等. 读者可以利用这些软件工具对所学的模型和计算方法进行计算和检验.

在我国, 运筹学基础一类的图书拥有大量读者, 这类图书的累计销售量有的已达几十万册. 但国内目前这类图书只有很少几种. 2006 年人民邮电出版社图灵公司邀我们翻译新出版的《运筹学导论》(第 8 版), 这体现了出版社对发展我国运筹学的重视. 由于书的篇幅宏大, 有 1 000 多页, 中译本分成上下两册出版. 上册包括原书中属于基础部分的 12 章, 下册包括属于提高部分的 12 章. 每册均可用作一个学期的教材. 本书第 2, 3, 4, 13 章及附录 A 由薛毅教授翻译, 第 5 章由侯思祥教授翻译, 第 6, 7, 8, 16, 20 章及附录 C 由朱建明博士翻译, 其余各章及附录 B, D, E 由刘德刚博士翻译, 由我校阅. 中译本难免有疏漏和翻译不妥之处, 敬请读者给予指正.

韩继业

2007 年 5 月于中科院

前　言

本书第 8 版对内容作了很多的修订，在教材的编排上突出反映运筹学中的应用问题和计算方法.

- 第 2 章通过城市规划、货币套利交易、投资、生产计划、混合配比、排序以及下料等实际问题的应用，主要介绍了线性规划的建模. 新增加的习题也涉及从水质管理、交通控制到军事领域等多个运筹问题.
- 第 3 章以一种简单和直接的方式介绍了一般性的线性规划灵敏度分析，包括对偶价格和简约费用 (reduced cost)，作为单纯性表计算部分的直接扩充.
- 本版的第 4 章主要是基于对偶性进行线性规划最优化后的分析.
- 针对旅行商问题 (Traveling Salesperson Problem, TSP)，介绍了一个基于 Excel 的组合式最近邻点反向启发式算法.
- 新增的第 17 章扩充了马尔可夫链的处理方法.
- 在全新的第 24 章里，详细介绍了 15 个实际应用案例. 对这些案例的分析通常涉及多种 OR 技术 (例如启发式算法和线性规划，或者线性整数规划和排队论)，用来进行建模、数据采集以及问题的求解计算等. 这些应用问题在相关的各章里都有引用，让读者能够充分了解实际生活中如何运用运筹学技术.
- 新的附录 E 收录了按照章节排列的约 50 个小型实用问题的例子.
- 本书还布置了 1 000 多个节后习题，其中题前标有星号 (*) 的表示附录 C 给出了相应的答案.①
- 每章开头都有*本章导读*，帮助读者了解教材内容，有效利用附带的软件程序.
- 把教材与软件相结合可以让读者对需要深入介绍的概念进行实际检验：
  1. 全书都用到了 Excel 程序，包括动态规划、旅行商问题、库存问题、层次分析法、贝叶斯概率、“电子化”统计表、排队问题、模拟、马尔可夫链以及非线性规划等. 一些程序中的交互式用户输入功能有助于对相应方法的更好理解.
  2. 对 Excel 规划求解程序的使用扩展到了全书，特别用在线性规划、网络规划、整数规划和非线性规划问题.
  3. AMPL® 是一种强大的商业化建模语言，本书将 AMPL 结合在大量的例题中，这些例子涉及线性、网络、整数和非线性规划问题. 附录 A 给出了 AMPL 的语句规则以及本书例题中所引用的语言素材.
  4. 本书中，TORA 仍然充当教学软件的重要角色.
- 所有与计算机相关的材料都相对独立，有的作为单独的章节，有的按照标题

① 原书附录 C 拆分到上、下两册. 本书将不包含习题答案，读者可到图灵网站 (www.turingbook.com) 免费注册下载. —— 编者注

优化方法处于非常重要的地位.

运筹学的非凡价值使得许多国家的大学里的运筹学系、管理科学系、经济学系、工业工程系、系统科学系、数学系、计算科学系等早已开设了关于运筹学及其一些分支学科的课程. 我国的情况也大致如此. 从适应于不同院系专业的学生学习运筹学的角度来考虑, 一本好的运筹学基础的教科书十分重要. 在国外关于运筹学基础的诸多教材中, Hamdy A. Taha 著的*Operations Research, An Introduction* 是非常优秀的一本. Taha 是美国阿肯色大学工业学院工业工程教授, 世界知名运筹学家. 他著的*Operations Research, An Introduction, Preliminary edition* 出版于 1968 年. 此书经过多次修改与扩增, 今年已出版了第 8 版. 该书被世界上多所大学采用为运筹学基础教材, 已有西班牙文、日文、俄文、土耳其文、印尼文等多种译本出版. 它有 3 大重要特色. 一是内容广泛、取材得当. 连同附带的光盘, 共有 24 章和 5 个附录, 内容涉及线性规划、运输问题、网络问题、目标规划、整数规划、动态规划、库存问题、非线性规划等确定性运筹分支, 以及随机动态规划、随机库存问题、排队系统、马尔可夫决策过程、决策分析、对策论、模拟问题、预报问题等随机性运筹分支. 这些内容覆盖了迄今运筹学所研究的大部分重要问题. 该书在取材上首先是重视对上述运筹问题的基本知识的讲解, 但对某些问题也包括了较高深的内容, 以满足不同读者的需要. 二是突出实用性. 书中各章总是通过若干实际问题的求解来引导出所要讲的运筹问题的数学模型. 这既凸显出这些运筹问题的实际背景, 也使读者可学到如何进行建模. 第 24 章详细地介绍了 15 个实际应用案例, 运用了多种运筹学技术进行建模、数据采集以及求解计算等. 附录 E 还收录了近 50 个应用例子. 作者精心收集和分析的这些实例来源于工业、商业、金融、社会、体育、娱乐等许多行业, 是很好的运筹学教学资料. 三是计算方法与软件相结合. 全书大量使用教学辅助软件 TORA、电子表格程序 Excel 及 AMPL 等. 读者可以利用这些软件工具对所学的模型和计算方法进行计算和检验.

在我国, 运筹学基础一类的图书拥有大量读者, 这类图书的累计销售量有的已达几十万册. 但国内目前这类图书只有很少几种. 2006 年人民邮电出版社图灵公司邀我们翻译新出版的《运筹学导论》(第 8 版), 这体现了出版社对发展我国运筹学的重视. 由于书的篇幅宏大, 有 1 000 多页, 中译本分成上下两册出版. 上册包括原书中属于基础部分的 12 章, 下册包括属于提高部分的 12 章. 每册均可用作一个学期的教材. 本书第 2, 3, 4, 13 章及附录 A 由薛毅教授翻译, 第 5 章由侯思祥教授翻译, 第 6, 7, 8, 16, 20 章及附录 C 由朱建明博士翻译, 其余各章及附录 B, D, E 由刘德刚博士翻译, 由我校阅. 中译本难免有疏漏和翻译不妥之处, 敬请读者给予指正.

韩继业

2007 年 5 月于中科院

目录

第 1 章　什么是运筹学

本章导读　最早期的较为正式的运筹学 (Operations Research, OR) 活动出现在第二次世界大战时期, 当时有一批英国的科学家着手研究如何利用科学方法进行决策, 以最佳地利用战时的资源. 战后, 人们对军事作战中提出来的这些运筹学思想进行了改进, 使之用于民用领域以提高工作效率和生产力.

本章将让你熟悉运筹学的基本术语, 包括数学建模、可行解、最优化和迭代运算等基本概念. 你将会了解到, 对问题做出正确的定义是运用运筹学最重要 (也是最困难) 的一步. 本章还强调, 虽然数学建模是运筹学最基本的工作, 但在最终决策时还必须考虑到一些无形的因素 (不能定量化的因素, 如人的行为). 随着对本书的学习, 你会接触到各式各样的应用实例, 有解题的例子, 也有各章的习题. 特别是第 24 章 (在下册), 全部是精心编制的案例分析. 各章内容均与这些案例互相配合, 以充分展现运筹学在实际中的运用情况.

1.1　运筹学模型

设想你有一项工作任务, 需要 5 周完成, 其间要往返于 Fayetteville (FYV) 与 Denver (DEN) 之间. 每个星期一你都要乘飞机从 Fayetteville 出发, 星期三返回. 普通的往返机票价格是 400 美元, 但如果机票往返期间内跨越周末的话, 则可以享受 20% 的票价折扣. 不论去程还是回程, 一张单程机票的价格都是普通往返机票的 75%. 那么, 你应该如何购买这 5 周期间的机票呢?

可以把这个例子看作是一个决策问题, 要求解这个问题需要回答 3 个提问:

(1) 都有哪些可能的**决策方案**?

(2) 是在什么**限制条件**下作出这个决策的?

(3) 评价这些方案的**目标评判标准**是什么?

考虑 3 种可能的决策方案:

(1) 购买 5 张普通的 FYV-DEN-FYV 往返机票, 每周星期一出发, 星期三返回.

(2) 购买 1 张 FYV-DEN 的单程机票和 4 张跨越周末的 DEN-FYV-DEN 往返机票, 再买 1 张 DEN-FYV 单程机票.

(3) 先购买 1 张第一周星期一出发、最后一周星期三返程的 FYV-DEN-FYV 往返票, 再买 4 张跨周末的 DEN-FYV-DEN 往返机票. 这一方案中所有机票都至少跨越一个周末.

对这些方案的限制条件都是, 你必须要每周星期一从 FYV 出发, 并在本周的星期三返回.

评价所提出的各种方案的一个明显的目标评判标准就是购买这些机票的总费用, 花费最少的方案最佳. 针对上述 3 种方案, 我们有

方案 1 的费用 $= 5 \times 400 = \$2\,000$

方案 2 的费用 $= 0.75 \times 400 + 4 \times (0.8 \times 400) + 0.75 \times 400 = \$1\,880$

方案 3 的费用 $= 5 \times (0.8 \times 400) = \mathbf{\$1\,600}$

因此, 你应该选择方案 3.

虽然上述例子表明了运筹学模型的 3 个主要的构成部分, 即备选方案、目标评判标准和约束条件, 但在对每个部分如何进行详细构造时会遇到各种各样的情况. 为了说明这一点, 我们考虑用一段长度为 L 英寸的电线来围成一个矩形, 要让这个矩形的面积最大, 它的长和宽各应该取多少呢?

和购买机票的例子不同的是, 现在这个例子中的可能方案数不是有限个, 因为矩形的长和宽可能有无限多个值. 为了把这个问题用公式表示出来, 我们把长和宽定义成两个连续的 (代数) 变量, 用来标识该问题的所有可能方案.

令 $w =$ 用英寸表示的矩形长; $h =$ 用英寸表示的矩形宽.

根据这些定义, 问题的限制条件可以叙述为

(1) 矩形长 + 矩形宽 = 电线长度的一半;　　(2) 长和宽不能为负值.

这些限制条件可用代数形式表示成

(1) $2(w + h) = L$;　　(2) $w \geqslant 0, h \geqslant 0$.

现在剩下的部分就是问题的目标了, 即让矩形的面积最大. 令 z 为该矩形的面积, 则整个模型就变成了

$$\begin{aligned} \max\ & z = wh \\ \text{s.t.}\ & 2(w+h) = L, \quad w,\ h \geqslant 0 \end{aligned}$$

这个模型的最优解为 $w = h = \frac{L}{4}$, 它要求构造的是一个正方形.

基于前面的两个例子, 一般的运筹学模型可以组织成下面的通用格式:

max 或 min	**目标函数**
s.t.	**约束条件**

一个模型的解如果满足所有的约束条件, 则称它是**可行的** (feasible); 如果既是可行的, 且又取得了目标函数的最佳 (最大或者最小) 值, 则称它是**最优的** (optimal). 在购机票的例子中, 该问题提出了 3 个可行的方案, 第 3 个方案得到了最优解. 而在构造矩形的例子中, 可行方案必须满足条件 $w + h = \frac{L}{2}$, w 和 h 取非负值. 这样就产生了无穷多个可行解, 和购机票问题不同的是, 这一最优解是通过适当的数学工具求出来的 (在这个例子中, 我们用了微积分方法).

虽然运筹学模型是要在一组约束条件下, 使得某一具体的目标评判标准达到"最优", 但它所得出的解的质量取决于模型对实际问题刻画的完全性. 以购机票问题为例, 假如我们不能够找出购买机票的所有方案的话, 那么所得到的解只相对于所选模型是最优的. 具体说, 假如方案 3 没有包括在模型中, 那么所得到的"最优"解就是要用 1 880 美元来购买这些机票, 这只是一个**次最优** (suboptimal) 解. 我们的结论是, 一个模型的"最优"解只是对这个模型是最好的, 只有当这个模型恰当地表达了实际问题时, 它的解对于实际问题才是最优的.

习题 1.1A

1. 在购买机票例子中, 找出第 4 种可能的方案.

2. 在构造矩形例子中, 找出两个可行解, 并计算哪一个更好.

3. 求出构造矩形问题的最优解. (提示: 利用约束方程表示出单变量的目标函数, 然后用微积分求解.)

4. Amy、Jim、John 和 Kelly 正站在一条河的东岸, 想用木筏划到河的西岸去. 木筏每次最多能坐两个人, Amy 身体最棒, 她能在 1 分钟内划过河, 而 Jim、John 和 Kelly 分别需要用 2 分钟、5 分钟和 10 分钟划过对岸. 假如有两个人在木筏上, 过河时间按照较慢的人算. 目标是要用尽可能短的时间让这 4 个人都过到河的对岸.

(a) 找出至少两个可行的过河方案 (记住, 木筏是唯一的交通方式, 而且不允许放空船).

(b) 定义评价各方案的评判标准.

***(c)** 把 4 个人都渡到对岸的最短时间是多少?

***5.** 在棒球比赛中, Jim 当投手, Joe 是击球队员. 假设 Jim 可随机投出快球或曲线球. 如果 Joe 能够正确判断出一个曲线球的话, 他就能保持 0.5 的平均击中率; 否则假如 Jim 投出一个曲线球, 而 Joe 按照快球来准备的话, 他的平均击中率则下降到 0.2. 另一方面, 如果 Joe 正确地判断出快球, 他能达到 0.3 的平均击中率; 否则, 他的击中率仅有 0.1.

(a) 针对上述情况, 找出可能的方案.

(b) 求该问题的目标函数, 并讨论它和我们熟悉的对某一评判标准的最优化 (最大化或最小化) 有什么不同.

6. 在建造一幢房子时, 共有 6 根 24 英尺长的地板龙骨, 每根都必须切割成 23 英尺长的成料, 切割龙骨的操作步骤如下:

操作步骤	所需时间 (秒)
(1) 把龙骨放置到锯床上	15
(2) 量出所需的长度 (23 英尺)	5
(3) 标画出圆锯切割线	5
(4) 把龙骨切割成所需的长度	20
(5) 把切好的龙骨堆放到指定的区域	20

这项作业需要 3 名工人操作: 两名装料工必须同时操作第 1 步、第 2 步和第 5 步, 一名切割工负责第 3 步和第 4 步操作. 共有两对锯床, 待锯的龙骨放在上面准备切割, 每对锯床最多可并排放置 3 根龙骨. 请给出切割这 6 根龙骨的一个好的工序安排方案.

1.2 运筹学模型的求解

在 OR 中, 我们并没有一种万能的技术能求解出实践中可能出现的所有数学模型. 恰恰相反, 数学模型类型的多样性和复杂程度的各异性决定了求解方法迥异的特性. 例如在 1.1 节中, 为求出购票问题的解, 只要对各方案按照机票总费用排个顺序就行了; 而对于构造矩形问题的求解, 就需要利用微积分来确定最大的面积.

线性规划(linear programming) 是最常用的 OR 技术, 专门用于带有线性目标函数和约束函数的模型. 其他方法还有**整数规划**(integer programming)(变量取整数值)、**动态规划**(dynamic programming)(初始模型可分解成多个较小的子问题)、**网络规划**(network programming)(问题可以刻画成一个网络), 以及**非线性规划** (nonlinear programming)(模型的函数是非线性的). 还有许许多多的其他运筹学方法.

大多数运筹学技术的一个特性是, 问题的解常常不是通过某种 (像解析式一样的) 闭形式 (closed form) 得到的, 而是利用某些**算法** (algorithm) 求出来的. 算法提供一些固定的计算规则, 利用它反复对问题进行计算, 每次重复计算 [称为**迭代** (iteration)] 所得到的解都向最优解逐步靠近. 由于每次迭代的计算往往都是类似的, 计算量又大, 这些算法必须在计算机上运行.

有些数学模型可能非常复杂, 已有的最优化算法都无法求出最优解来. 在这种情况下, 可能必须放弃寻找*最优解*, 这就需要利用某些**启发式算法**或某些*经验方法*, 找到一个*较好的解*.

1.3 排队模型和模拟模型

排队和模拟用于研究等待队列, 它们不属于最优化技术, 而是用来度量等待队列的性能, 例如队列中的平均等待时间、服务的平均等待时间, 以及服务设施的利用率等.

排队模型利用概率论和随机模型对等待队列进行分析, 模拟则是通过模仿实际系统的行为, 来估计这些性能指标. 从某种意义上, 模拟可以被认为是观察实际系统的一种最好的方法. 排队和模拟之间的主要差别在于, 排队模型是纯数学的, 因此不得不服从于具体的假设, 这就限制了它们的应用范围, 而模拟却非常灵活, 可用来分析任何实际的排队情形.

模拟的使用也并非没有缺点. 建立模拟模型的过程既费时又费力, 此外, 即使在最快的计算机上运行模拟模型, 通常也很慢.

1.4 建模的艺术

1.1 节中所建的示例模型是对实际情况的真实表达, 这在运筹学中是很少出现的, 因为大多数应用问题中, 通常都涉及 (不同程度的) 各种近似. 图 1.1 表示运筹学建模过程中表现出来的抽象水平. 我们把注意力放在控制实际系统行为的主要变量上, 从真实情况中抽象出假定的实际系统. 这样的模型用某种可处理的方法, 表达出代表这个假定的实际系统行为的数学模型.

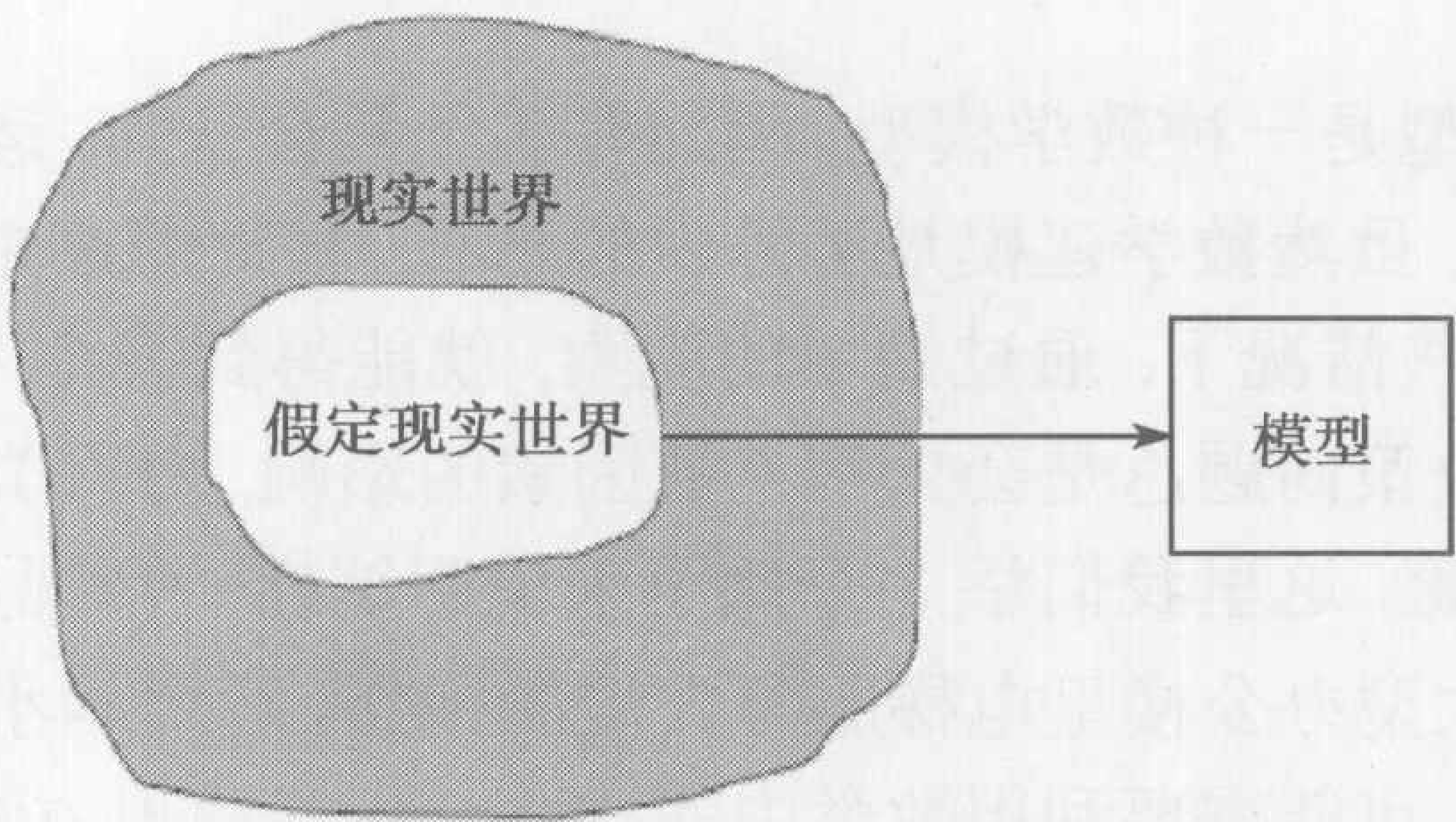


图 1.1 建模中的抽象水平

为了说明建模中的不同抽象水平, 我们以 Tyko 制造公司为例. 该公司生产各种塑料容器制品, 当一份生产订单下达到生产部门时, 必要的原材料要从公司的库存获得或从外面购买. 一旦完成了批量生产以后, 销售部门负责向客户分销这些产品.

在对 Tyko 公司的情况进行分析时, 一个自然的问题是要决定生产批量的大小. 那么, 如何用模型来表达这个问题呢?

考察整个系统后我们发现, 有许多变量都可以直接用来表示生产水平, 下面是按照部门分类的一部分变量:

(1) 生产部门: 用现有机器、工人工作时间、半成品库存量, 以及质量控制标准表示的生产能力.

(2) 原材料部门: 现有原材料库存量、采购供货安排、库存限量.

(3) 销售部门: 销售预测、分销网能力、广告促销能力和竞争效果.

在这些变量中, 每一个都影响着 Tyko 公司的生产量水平, 要想在这些变量与产量水平之间建立起明确的函数关系的确并非易事.

第一个抽象水平需要定义出假定实际系统的边界. 通过仔细分析, 我们可以用下面两个主要变量来近似描述实际系统:

(1) 生产率; (2) 消费率.

计算生产率要用到生产能力、质量控制标准、现有原材料等一些变量, 消费率则可以从与销售部门有关的变量中算出. 本质上, 从现实世界到假定现实世界的简化, 是通过把多个现实世界变量“简化”成某个单一的假定现实世界变量来实现的.

这样, 从假定的现实世界抽象出一个模型就比较容易了. 从生产率和消费率中, 就可以建立起库存剩余或不足的度量, 然后就可以建立起抽象出来的模型, 用来平衡库存过剩或库存短缺所引起的冲突成本, 使得库存的总费用达到最少.

1.5　仅有数学是不够的

由于运筹学模型是一种数学模型, 因此有人经常会认为, 运筹学的运用总是要根植于数学的分析. 虽然数学建模是运筹学的基石, 但我们首先还是应该运用一些简单的方法. 在一些情况下, 通过简单的观察, 就能得到某个“常识性的”解. 其实, 由于大多数的决策问题总是会受到人的因素的影响, 对人类心理的研究可能就成为解决问题的关键. 这里我们举 3 个例子来说明这样一个观点:

(1) 针对某个大型办公楼里电梯服务太慢的抱怨, 运筹学小组一开始觉得这是一个等待队列问题, 可能需要利用数学中的排队分析或模拟方法来解决. 在对提出这些抱怨的人们的行为进行研究以后, 小组里的心理学专家提议, 在电梯口安装一些落地镜子. 不可思议的是, 这些抱怨随之消失了, 因为人们在等待电梯时, 只顾着自己照镜子了.

(2) 一个美加联合专家小组用排队论对英国某大机场值机柜台的实际情况进行研究和分析. 建议在适当位置放置一些指示牌, 以便让离登机时间不足 20 分钟的紧急旅客可以直接排到队首, 申请即刻办理登记手续. 但这个解决办法并不奏效, 因为大部分旅客是英国人, 他们“习惯于非常严格的排队行为”, 因此不愿意插到其他排队旅客的前面.

(3) 在某钢铁厂, 先用铁矿石炼出钢锭, 然后用来制造钢条和钢梁. 管理人员注意到, 从钢锭生产到运送到下一个生产环节 (制成最终产品) 之间, 要耽误很长的时间. 在理想情况下, 为了降低重新加热的成本, 应该在钢锭离开熔炉后马上开始钢梁的制造. 一开始, 这个问题被看成是一个生产线均衡问题, 为解决这个问题, 要么减少钢锭的产量, 要么提高制造过程的能力. 运筹学小组利用了简单的图表, 把每天三班期间炼炉的产量进行汇总, 他们发现, 即使第三班从晚上 11 点就开始, 但大部分的钢锭还是在早上 2 点到 7 点之间生产出来的. 进一步调查还发现, 第三班工人愿意在刚开始接班的时候多休息一会儿, 因此造成了早班产量的下降. 最终问题的解决方案是, 让整个第三班期间的钢锭产量“均衡”下来.

从上述这些例子中我们可以得出 3 个结论:

(1) 在着手构造复杂的数学模型之前, 运筹学专家应该探讨是否能用“突破常

规的”思路来解决问题. 通过安装镜子来解决电梯问题是利用了人们的心理而不是用数学建模的方法, 比起数学模型可能给出的建议来, 这个方法更加简单, 也更加经济. 这也是为什么运筹学小组中通常要有来自非数学领域的“外来”专家 (如电梯问题例子中的心理学专家) 的原因. 第二次世界大战期间, 当时英国的第一个运筹学小组就提出并贯彻了这一观点.

(2) 解决问题要考虑的关键是人而不是技术. 在解决方案中不考虑人的行为, 是必定要失败的. 尽管对英国机场问题的数学求解方案可能是完全正确的, 但事实上专家小组并不了解美国人和英国人之间的文化差异 (美国人和加拿大人往往就不那么讲究礼数), 因此导致所建议的解决方案不能实现.

(3) 在运用运筹学解决某个问题之前, 决不应先入为主地认定要使用某个具体的数学工具, 要先弄清楚使用这个数学工具是否合理. 例如, 不能因为线性规划是一项成功的技术, 就倾向于用它对“任何”情况进行建模. 这样一种方法会导致所建立的数学模型偏离实际, 因此必须要首先对已有的资料加以分析, 尽可能采用最简单的技术 (例如, 平均值、图表等), 目的是要抓住问题的源头. 一旦对问题做出了准确的定义, 就可以决定使用最合适的求解工具[①]. 在钢铁厂问题中, 钢锭生产的简单图表对于弄清情况是最需要的.

1.6 运用运筹学的几个步骤

运用运筹学解决实际问题需要*团队精神*, 需要 OR 分析人员和客户并肩努力. 运筹学分析人员在建模方面的专业技能必须要和客户的经验与合作配合起来, 要针对客户的问题开展研究.

作为决策的工具, 运筹学既是一门科学, 更是一门艺术. 说它是一门科学, 因为它体现了数学技术的优点; 说它是一门艺术, 因为成功地实现数学模型求解的每一个步骤, 大都取决于运筹学小组的创造性和经验. Willemain(1994) 认为, “有效的 (OR) 实际应用不仅仅需要定量分析的能力, 还需要对所用技术的判断力 (比如在什么时候以及如何使用某项具体的技术), 需要与人沟通以及组织生存方面的技巧.”

对于这些无形的因素, 很难规定具体的 (类似于数学模型的精确原理所叙述的) 做法, 但我们还是能够对如何在实际中运用运筹学提出一些通行的指导原则.

在实际中运用运筹学的主要步骤包括:

(1) 问题定义; (2) 模型构造; (3) 模型求解;

① 在弄清某个具体的数学模型是否合理之前就决定采用它, 就好比是“本末倒置”. 这让我想起了一个故事, 说的是有位经常坐飞机的旅客, 他很偏执地想象飞机上可能有恐怖袭击炸弹. 他计算了这种事件发生的概率, 虽然这个概率非常小, 但还是不能消除他的忧虑. 从那儿以后, 他总是在行李箱里自己携带一枚炸弹上飞机, 因为根据他的计算, 飞机上有两枚炸弹的概率几乎为零.

(4) 模型验证; (5) 解决方案实施.

其中第 3 步*模型求解*最明确, 也是在运用运筹学解决实际问题中最容易实现的, 因为这一步骤主要是针对特定的数学模型, 而其他步骤的实现则更多地是一门艺术, 而不是一种理论.

问题定义涉及确定所研究问题的范围, 这项工作需要全体 OR 小组共同来做, 目的是找出这个决策问题的 3 个要素: (1) 描述可能的决策方案; (2) 确定问题的目标; (3) 指出所建模的系统运行中的限制条件.

模型构造要求把问题的定义转化成数学关系. 如果产生的模型恰好是某种标准的数学模型, 比如线性规划, 我们通常就可以利用已有的算法去求出一个解. 反之, 假如数学关系太复杂求不出解析解来, 运筹学小组就可能要简化模型并使用启发式方法, 或者考虑利用模拟的方法. 在某些情况下, 就像第 24 章的案例分析中展示的那样, 还需要将数学的、模拟的以及启发式的模型相结合, 来求解一个决策问题.

模型求解是所有 OR 步骤中最简单的, 因为它只需要利用一些成熟的最优化算法. 模型求解步骤中一个最重要的内容是进行*灵敏度分析*, 为的是要了解当模型参数发生某些改变时, 最优解会有怎样的表现. 尤其当模型的参数不能准确估计时, 更需要做灵敏度分析. 在这些情况下, 在所估计参数的邻域内研究最优解的表现是非常重要的.

模型验证是为了检查所提出的模型是否体现了真正的意图, 也就是说, 模型是否能充分预知所研究系统的行为. 从一开始, OR 小组就应该确信, 这个模型的输出不会出现"意外情形", 换句话说, 得出的解是合理的吗? 这些结果在直观上是可接受的吗? 正式地讲, 检查一个模型是否正确的一般方法, 是把模型的输出结果与历史的输出数据进行比较, 如果在相似的输入条件下, 模型合理地重现了过去的情形, 则这个模型就是正确的. 但一般情况下, 我们并不能保证, 未来的情形会继续重复过去的行为. 还有, 因为你的模型通常是基于对历史数据的仔细分析, 所提出的比较往往令人满意. 假如所提出的模型表达的是一个 (从未出现过的) 新系统, 没有现成的历史数据, 在这种情况下, 我们可以利用模拟作为一种独立的工具, 来检验数学模型的输出结果的正确性.

实施一个经过正确性检验的模型的解, 需要把得到的结果翻译成一些能让人明白的操作指令, 下达给所研究的系统的管理人员. 这项任务主要是由 OR 小组来做的.

1.7 关于本书

Morris(1967) 里写道: "讲授模型并不等同于讲授建模." 我在准备本书第 8 版时就已经注意到这句话的重要含义, 并努力通过在全书中收集一些实际的模型, 来

介绍 OR 中的建模艺术. 由于运筹学中计算的重要性, 本书还介绍了各种计算工具, 从教学辅助软件 TORA 到商业化的软件包 Excel、Excel 规划求解和 AMPL.

运筹学基础课应该在教给学生一些 OR 的数学基础的同时, 还要让学生了解运筹学的一些应用问题. 这会为运筹学的使用者提供某种信心, 假如我们的训练一味地集中在 OR 的理念和艺术方面的话, 这种信心通常就会丧失掉. 在建立了数学基础以后, 你就可以通过学习所发表的实际案例, 来提高你在 OR 建模艺术方面的能力. 为了在这方面为读者提供帮助, 第 24 章中收集了 15 个精心编写和详细分析的案例, 涉及了本书中介绍的大部分 OR 模型. 在附录 E 中, 还给出了大约 50 个实际生活中的应用实例. 在杂志和出版物中, 还可以找到更多的案例研究. 特别是由 INFORMS(美国运筹与管理科学学会) 出版的*Interfaces*, 介绍了各种各样的运筹学应用实例.

参 考 文 献

Altier, W.J., *The Thinking Manager's Toolbox: Effective Processes for Problem Solving and Decision Making*, Oxford University Press, New York, 1999.

Checkland, P., *Systems Thinking, System Practice*, Wiley, New York, 1999.

Evans, J., *Creative Thinking in the Decision and Management Sciences*, South-Western Publishing, Cincinnati, 1991.

Gass, S., "Model World: Danger, Beware the User as a Modeler," *Interfaces*, Vol.20, No.3, pp.60-64, 1990.

Morris, W., "On the Art of Modeling," *Management Science*, Vol.13, pp.B707-B717, 1967.

Paulos, J.A., *Innumeracy: Mathematical Illiteracy and its Consequences*, Hill and Wang, New York, 1988.

Singh, Simon, *Fermat's Enigma*, Walker, New York, 1997.

Willemain, T. R., "Insights on Modeling from a Dozen Experts," *Operations Research*, Vol.42, No.2, pp.213-222, 1994.

第2章 线性规划建模

本章导读 本章着重介绍线性规划 (linear programming, LP) 模型的建立与计算. 我们以二维变量问题的建模和图解开始, 尽管非常简单, 但它提供了理解线性规划基本概念的框架, 为第 3 章介绍的一般形式的单纯形算法奠定了基础. 为了阐明线性规划在实际生活中的使用, 我们对不同领域的一些应用问题建立了模型并求解, 这些问题包括: 城市规划、货币套汇、投资、生产计划和库存控制、汽油混合、人力规划和排序. 在计算方面, 本章使用两种不同类型的软件: (1) TORA, 它是一种完全菜单驱动和自文档化的说明程序, 通过交互式反馈, 帮助你理解线性规划的基本概念; (2) 基于电子表格的 Excel 规划求解和 AMPL 建模语言, 它们是为解决实际问题而设计的商业软件包.

2.1 节和 2.2 节的内容对于进一步理解本书中线性规划的内容是至关重要的. 你会发现 TORA 的交互式图解模式, 连同 2.2 节的内容, 对你特别有帮助. 2.3 节提出了各种形式的线性规划应用, 而且每种应用都配有相关的习题.

2.4 节介绍商业软件包 Excel 规划求解和 AMPL. 我们用它们求解了 2.3 节中的模型, 而且所有的程序代码均包含在文件夹 ch2Files[①]中. 以后的各章还介绍了其他一些规划求解与 AMPL 模型, 有关 AMPL 语法的详细描述将在附录 A 中予以介绍. 学习 AMPL 软件和规划求解的一种好方法就是, 亲手实践贯穿全书的众多模型, 并且适当地修改它们以便求解每节后面的问题. 将 AMPL 代码和附录 A 中的内容相互参照, 将使你的学习过程变得更容易.

我们刻意把 TORA、规划求解和 AMPL 的材料或者安排在各自的小节中, 或者安排在小标题为 TORA/Solver/AMPL 程序片段中, 其目的是尽量不打乱主要的文字叙述. 尽管如此, 依然鼓励你使用计算机完成每节后面的习题. 因为有时, 一个模型在求解之前看起来可能是“正确的”, 只有在试图求解的时候, 你才会发现某些公式需要修改.

本章包括 2 个实际应用、12 个例子、2 个规划求解模型、4 个 AMPL 模型、94 个节后习题和 4 个案例. 这些案例在本书 (下册) 的附录 E 中. AMPL/Excel/Solver/TORA 程序在下载资料的文件夹 ch2Files 中.

实际应用——Frontier 航空公司如何经济地购买燃料

在飞行航线的任何中途停留处都可以为飞机加燃油. 各地燃油的价格是不同

① 文件夹在下载资料中. —— 译者注

的, 在一个较廉价的地点加载额外的燃油 [称为机动加油量 (tankering)] 为随后的航程使用, 就可能节省燃油费用. 机动加油量的缺点是, 额外的重量会烧掉更多的燃油. 使用线性规划 (和启发式算法) 来确定最优的机动加油量, 来平衡额外烧掉燃油的费用与在购进燃油时节省的费用. 1981 年完成的这项研究, 导致的净节约资金大约为每年 350 000 美元. 第 24 章的案例 1 提供了研究的细节. 有趣的是, 随着近年来燃料价格的上涨, 许多航空公司现在正在使用基于线性规划的机动加油量软件购买燃料.

2.1 二维变量的线性规划模型

本节介绍二维变量线性规划的图解法. 尽管二维变量问题在实际中很少出现, 但它的解法为在第 3 章提出的一般单纯形法奠定了坚实的基础.

例 2.1-1 (Reddy Mikks 公司)

Reddy Mikks 使用 M1 和 M2 两种原料生产内、外墙涂料. 下表提供了问题的基本数据.

	每吨产品使用原料的吨数		日最大可
	外墙涂料	内墙涂料	用量 (吨)
原料 M1	6	4	24
原料 M2	1	2	6
每吨利润 ($1 000)	5	4	

一项市场调查指出：内墙涂料的日需求量不超过外墙涂料的日需要量加上 1 吨. 同时, 内墙涂料的最大日需求量是 2 吨.

Reddy Mikks 打算确定最优的 (最好的) 内、外墙涂料的产品混合, 使得日总利润达到最大.

线性规划模型, 与任何运筹学模型一样, 由 3 个基本部分组成.

(1) 寻求需要确定的决策**变量** (variable).

(2) 需要优化 (求极大或求极小) 的**目标函数** (objective).

(3) 解必须满足的**约束** (constraint).

合理地确定决策变量是模型建立过程中重要的第一步. 一旦决策变量确定后, 构造目标函数和约束函数的工作就变得非常简单了.

对于 Reddy Mikks 问题, 我们需要确定内、外墙涂料的日生产量. 因此, 模型的变量定义为

$$x_1 = \text{外墙涂料的日生产吨数}, \quad x_2 = \text{内墙涂料的日生产吨数}$$

首先构造目标函数. 注意, 公司打算极大化(也就是尽可能地增加) 两种涂料的日总利润. 已知每吨外、内墙涂料的利润分别是 5 和 4(单位：1 000 美元), 所以有

$$\text{外墙涂料的总利润} = 5x_1, \quad \text{内墙涂料的总利润} = 4x_2$$

令 z 表示日总利润, 则公司的目标是

$$\max\ z = 5x_1 + 4x_2$$

下一步, 构造限制原料用量和产品需求量的约束. 原料限制可表述为

$$\text{生产两种涂料的原料用量} \leqslant \text{最大原料可用量}$$

原料 M1 的日用量是每生产 1 吨外墙涂料用 6 吨, 每生产 1 吨内墙涂料用 4 吨, 因此,

$$\text{生产外墙涂料的原料 M1 用量} = 6x_1 \text{吨/天}$$
$$\text{生产内墙涂料的原料 M1 用量} = 4x_2 \text{吨/天}$$

所以,

$$\text{生产两种涂料的原料 M1 用量} = (6x_1 + 4x_2)\ \text{吨/天}$$

同理可得

$$\text{生产两种涂料的原料 M2 用量} = (1x_1 + 2x_2)\ \text{吨/天}$$

因为原料 M1 和 M2 的日可用量分别限制在 24 吨和 6 吨, 相应的限制给出如下:

$$6x_1 + 4x_2 \leqslant 24\ \ (\text{原料 M1}), \quad x_1 + 2x_2 \leqslant 6\ \ (\text{原料 M2})$$

第 1 个需求限制规定, 内墙涂料的日生产量超过外墙涂料的日生产量的部分 (即 $x_2 - x_1$) 不能超过 1 吨, 用数学语言描述, 即是

$$x_2 - x_1 \leqslant 1 \quad (\text{市场限制})$$

第 2 个需求限制规定, 内墙涂料的最大日需求量限制在 2 吨以内, 即

$$x_2 \leqslant 2 \quad (\text{需求限制})$$

一个隐含的 (或自然的) 限制条件是, 变量 x_1 和 x_2 不能出现负值. **非负性限制** (nonnegativity restriction), $x_1 \geqslant 0$, $x_2 \geqslant 0$, 表达了这一要求.

完整的 Reddy Mikks 模型是

$$\max\ z = 5x_1 + 4x_2$$

$$\begin{aligned}
\text{s.t.}\quad 6x_1 + 4x_2 &\leqslant 24 && (1)\\
x_1 + 2x_2 &\leqslant 6 && (2)\\
-x_1 + x_2 &\leqslant 1 && (3)\\
x_2 &\leqslant 2 && (4)\\
x_1,\ x_2 &\geqslant 0 && (5)
\end{aligned}$$

满足全部 5 个约束的 x_1 和 x_2 的任意值构成一个**可行解** (feasible solution). 否则, 其解是**不可行的** (infeasible). 例如, 解 $x_1 = 3$ 吨/天、$x_2 = 1$ 吨/天是可行的, 因为

它没有违反包括非负限制在内的任何约束. 为了核实这个结果, 将 $x_1 = 3, x_2 = 1$ 代入每个约束的左端. 在约束 (1) 中, 我们有 $6x_1 + 4x_2 = 6 \times 3 + 4 \times 1 = 22$, 它小于约束的右端项 24. 约束 (2) 至约束 (5) 也有类似的结论 (请验证!). 而 $x_1 = 4$ 和 $x_2 = 1$ 这个解是不可行的, 因为它不满足约束 (1), 也就是, $6 \times 4 + 4 \times 1 = 28$, 它大于右端项 24.

问题的目标是寻找到最好的可行解, 称为**最优解** (optimum), 使得总利润达到最大. 在寻找最优解之前, 我们需要知道 Reddy Mikks 问题到底有多少个可行解. 答案是, 正如我们将在 2.2 节图解法中看到的那样, 有"无穷多个", 因此采用枚举法求解此问题是不可能的. 这样, 就需要一个有步骤的方法, 能够让我们在有限步内找到最优解. 在 2.2 节介绍的图解法和将第 3 章介绍的一般代数化方法将解释如何能够做到这一点.

线性规划模型的性质 在例 2.1-1 中, 目标和约束均是线性的. 线性蕴涵着线性规划必须满足 3 条基本性质:

(1) 比例性 这条性质要求每个决策变量无论是在目标函数还是在约束中, 其贡献与决策变量的值*直接成比例*. 例如, 在 Reddy Mikks 模型中, 其量 $5x_1$ 和 $4x_2$ 是按照给定的比例常数, 每吨的单位利润 5 和 4, 分别给出了生产 x_1 和 x_2 吨外、内墙涂料的利润. 另一方面, 如果 Reddy Mikks 允许当销售量超过某些量时, 超过部分的给予某种程度的折扣, 则利润将不再与生产量 x_1 和 x_2 成比例, 此时, 利润函数变成非线性的.

(2) 可加性 这条性质要求所有变量在目标函数和约束中的总贡献等于每个变量各自贡献的直接和. 在 Reddy Mikks 模型中, 总利润等于两个各自利润分量的和. 然而, 如果两种产品的市场占有份额是*竞争的*, 即一种产品的销售量增加会影响到另一种产品销售, 则可加性不再满足, 此时模型不再是线性的.

(3) 确定性 线性规划模型中所有目标函数和约束的系数都是确定的. 这意味着它们是已知的常数——这在实际中很少出现, 因为实际中的数据更可能被表示成概率分布. 本质上, 线性规划的系数是概率分布平均值的近似. 如果这些分布的标准差充分小, 则这种近似是可接受的. 大标准差问题可直接地用随机线性规划算法 (见 19.2.3 节) 或者间接地用灵敏度分析 (见 3.6 节) 来求出它的最优解.

习题 2.1A

1. 对于 Reddy Mikks 模型, 构造下面约束的每一种情况, 并将它们表示成线性的左端项和常数的右端项:
   - *(a) 内墙涂料的日需求量比外墙涂料的日需求量*至少多* 1 吨.
   - (b) 原料 M2 的日使用量*最多*是 6 吨, *至少*是 3 吨.
   - *(c) 内墙涂料的需求量不能少于外墙涂料的需求量.
   - (d) 内、外墙涂料的最小生产量都是 3 吨.

*(e) 内墙涂料与内、外墙涂料的总产量之比不能超过 0.5.

2. 在下面 Reddy Mikks 模型的 (可行和不可行) 解之中, 确定最好的可行解:

(a) $x_1 = 1, x_2 = 4$　**(b)** $x_1 = 2, x_2 = 2$　**(c)** $x_1 = 3, x_2 = 1.5$

(d) $x_1 = 2, x_2 = 1$　**(e)** $x_1 = 2, x_2 = -1$

***3.** 对于 Reddy Mikks 模型的可行解 $x_1 = 2, x_2 = 2$, 确定原料 M1 和 M2 没有用完的量.

4. 假定 Reddy Mikks 采用按量打折销售的方法, 把它的外墙涂料卖给某个批发商. 如果批发商每日的购买量不超过 2 吨, 则 Reddy Mikks 每吨的利润是 5 000 美元, 否则是 4 500 美元. 给出目标函数的数学表达式, 得到的函数还是线性的吗?

2.2　线性规划的图解法

图解法的过程包括以下两步:

(1) 可行解空间的确定;

(2) 从可行解空间所有的可行点中确定最优解.

这个过程用两个例子来表明如何处理极大化和极小化目标函数.

2.2.1　极大化模型的解

例 2.2-1

这个例子求解例 2.1-1 中的 Reddy Mikks 模型.

第 1 步　可行解空间的确定.

首先, 我们说明如何表示对非负性约束 $x_1 \geqslant 0$ 和 $x_2 \geqslant 0$. 在图 2.1 中, 横坐标

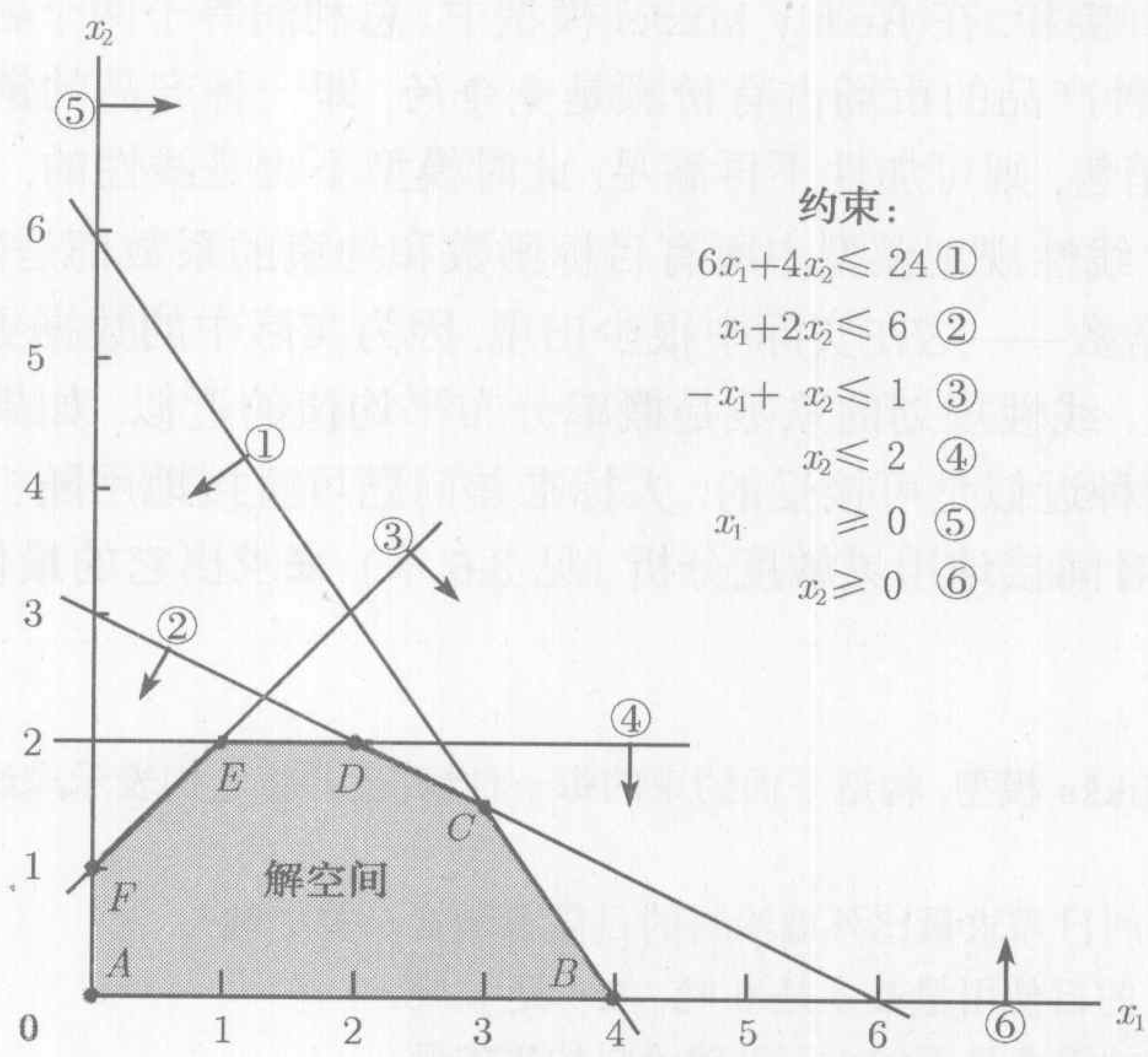


图 2.1　Reddy Mikks 模型的可行空间

x_1 和纵坐标 x_2 分别表示外、内墙涂料的变量. 因此, 变量的非负性限制解空间的区域在第一象限, 即 x_1 轴的上方和 x_2 轴的右方.

为说明其余的 4 个约束, 首先将每个不等式用等式方程替换, 然后, 确定方程上两个不同点, 画出相应的直线. 例如, 在用方程 $6x_1 + 4x_2 = 24$ 替换不等式 $6x_1+4x_2 \leqslant 24$ 后, 我们可以确定两个不同的点, 如首先令 $x_1 = 0$, 得到 $x_2 = \frac{24}{4} = 6$, 然后令 $x_2 = 0$, 得到 $x_1 = \frac{24}{6} = 4$. 因此, 可作出通过两点 $(0,6)$ 和 $(4,0)$ 的直线, 如图 2.1 的直线①所示.

下面考虑不等式的作用. 不等式的作用就是将 (x_1, x_2) 平面分成两个半空间, 每个半空间位于所画直线的一侧. 这两个半空间中只有一个满足不等式. 为了确定正确的一侧, 选择 $(0,0)$ 作为*参考点*(referrence point). 如果它满足不等式, 则参考点所在的一侧是可行的半空间, 否则另一侧是可行的. 我们用参考点 $(0,0)$ 来解释约束 $6x_1 + 4x_2 \leqslant 24$. 因为 $6\times 0+4\times 0=0$ 是小于 24 的, 所以代表该不等式的半空间包含原点 (见图 2.1 中箭头所指的方向).

为计算方便起见, 选择 $(0,0)$ 作为参考点, 除非直线恰好通过原点, 此时可以使用其他的点作为参考点. 例如, 我们用的参考点是 $(6,0)$, 第一个约束的左端是 $6\times 6+4\times 0=36$, 大于它的右端 24, 这意味着 $(6,0)$ 位于不等式 $6x_1 + 4x_2 \leqslant 24$ 中不可行的一侧. 这个结论与基于参考点 $(0,0)$ 得到的结论是一致的.

对于模型的所有约束, 应用参考点过程产生了由图 2.1 所示的约束 (请验证!). 问题的**可行解空间** (feasible solution space) 表示为第一象限内所有约束同时满足的区域. 在图 2.1 中, 在区域 $ABCDEF$ 边界内或边界上的任何点均是可行空间的一部分. 这个区域外的所有点都是不可行的.

TORA 程序

菜单驱动的 TORA 软件的图解线性规划模块, 将会帮助你加深对如何绘制线性规划约束图形的理解. 从 MAIN 菜单 (主菜单) 中选择 Linear Programming (线性规划). 在输入完模型之后, 从 SOLVE/MODIFY (求解/修改) 中选择 Solve (求解)⇒ Graphical (图形的). 在输出屏幕中, 你可以通过作交互式的实验, 逐一地对每个约束绘出图形, 这样你就能够看到每个约束是如何影响解空间的.

第 2 步 *最优解的确定*.

图 2.1 的可行空间是由连接点 A, B, C, D, E, F 的线段描述的. 在空间 $ABCDEF$ 边界内或边界上的任意一点均是可行的. 因为可行空间 $ABCDEF$ 由*无穷*个点构成, 我们需要一个系统化的过程来确定最优解.

求最优解需要确认利润函数 $z = 5x_1 + 4x_2$ 增加的方向 (这里我们是极大化 z). 我们可以通过指定对 z 的任意增加值来完成. 例如, 令 $z = 10$ 和 $z = 15$, 这等价于

画出的两条直线 $5x_1 + 4x_2 = 10$ 和 $5x_1 + 4x_2 = 15$. 因此, z 增加的方向由图 2.2 所示. 最优点出现在 C 处, 该点位于解空间中, 超过这一点进一步增加 z 值的任何点均在 $ABCDEF$ 的边界以外.

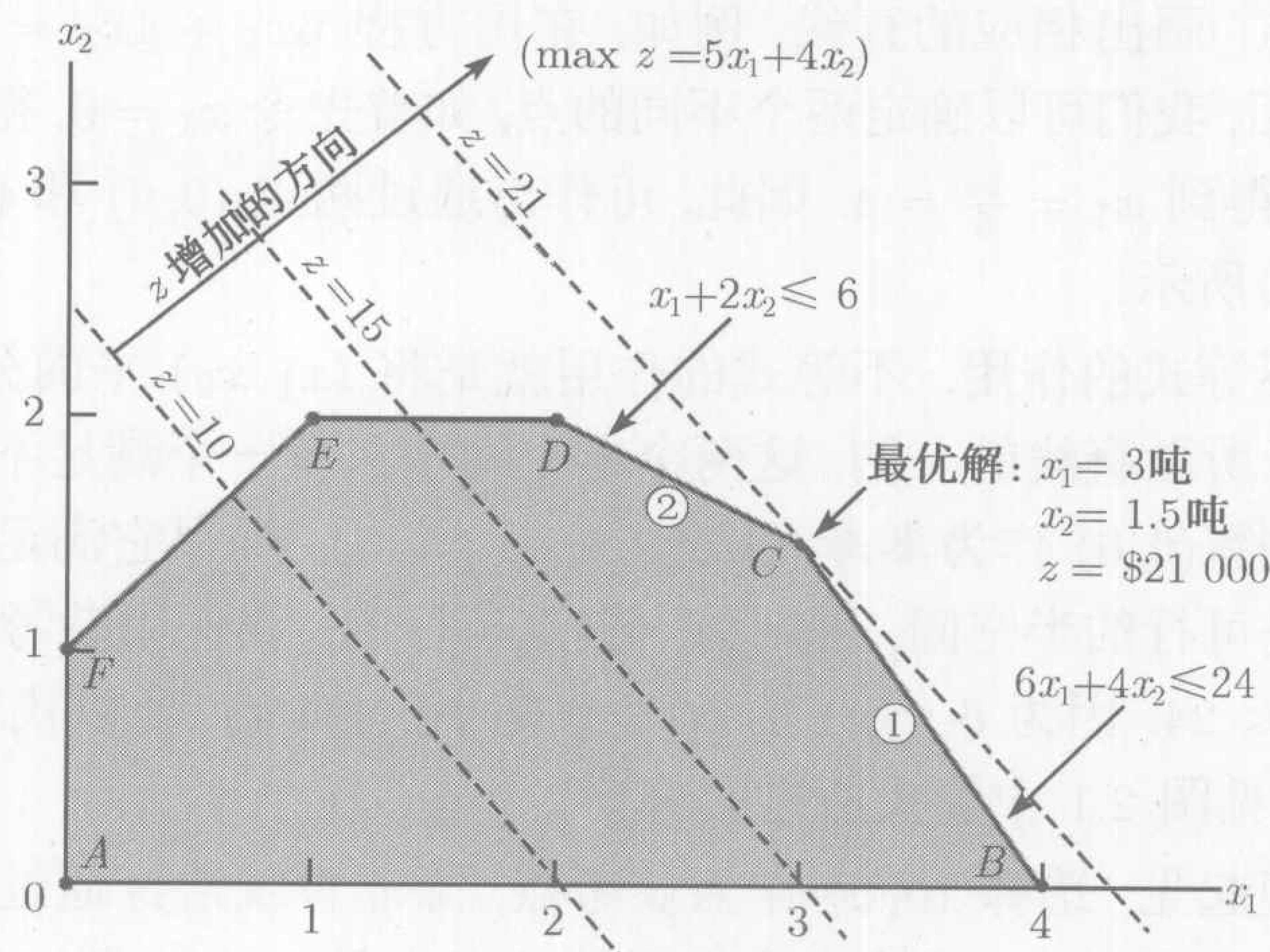


图 2.2　Reddy Mikks 模型的最优解

最优点 C 处 x_1 和 x_2 的值由直线①和直线②所构成方程的解确定, 即

$$6x_1 + 4x_2 = 24$$

$$x_1 + 2x_2 = 6$$

其解是 $x_1 = 3$ 和 $x_2 = 1.5$, 相应的 $z = 5 \times 3 + 4 \times 1.5 = 21$. 这个解要求每日生产 3 吨外墙涂料和 1.5 吨内墙涂料, 相应的日利润是 21 000 美元.

线性规划最优解的一个重要特征是, 最优解总是发生在解空间的**角点** (corner point)(这里是两条直线的交点) 处. 即使目标函数恰好平行于某个约束, 这一结论也是正确的. 例如, 如果目标函数是 $z = 6x_1 + 4x_2$, 它平行于约束 (1), 我们总可以说最优解或者出现在角点 B, 或者出现在角点 C. 实际上, 线段 BC 上任意点都是最优点 (见例 3.5-2), 但这里重要的事实是, 线段 BC 完全由角点 B 和 C 确定.

TORA 程序

可以用 TORA 观察到, 最优解总是在某个角点处得到. 从输出屏幕中, 可以单击 View/Modify Input Data (观察/修改输入数据) 修改目标函数的系数, 再用图解法求解该问题. 可以用下列目标函数检验上述概念:

(a) $z = 5x_1 + x_2$　(b) $z = 5x_1 + 4x_2$　(c) $z = x_1 + 3x_2$

(d) $z = -x_1 + 2x_2$　(e) $z = -2x_1 + x_2$　(f) $z = -x_1 - x_2$

观察到线性规划的最优解总是发生在角点处, 这意味着最优解能够简单地通过枚举所有角点来找到, 正如下表所示:

角点	(x_1, x_2)	z
A	$(0,0)$	0
B	$(4,0)$	20
C	**(3,1.5)**	**21(最优解)**
D	$(2,2)$	18
E	$(1,2)$	13
F	$(0,1)$	4

随着约束和变量个数的增加, 角点的个数也随之增加, 所提出的枚举法由于计算量太大而变得很繁琐. 尽管如此, 从确定线性规划最优解的角度来说, 上述思想表明, 包含*无穷*个解的解空间 $ABCDEF$ 实际上可以用有限个有希望成为解的点——角点 A, B, C, D, E, F 来替代. 这个结果是构造一般代数式算法也即*单纯形方法*的关键, 我们将在第 3 章进一步研究这种方法.

习题 2.2A

1. 已知 $x_1, x_2 \geqslant 0$, 确定下面每个独立约束的可行空间.

*(a) $-3x_1 + x_2 \leqslant 6$ (b) $x_1 - 2x_2 \geqslant 5$ (c) $2x_1 - 3x_2 \leqslant 12$

*(d) $x_1 - x_2 \leqslant 0$ (e) $-x_1 + x_2 \geqslant 0$

2. 确定下列情况中 z 的增加方向:

*(a) max $z = x_1 - x_2$ (b) max $z = -5x_1 - 6x_2$

(c) max $z = -x_1 + 2x_2$ *(d) max $z = -3x_1 + x_2$

3. 单独改变下面每个值, 确定 Reddy Mikks 模型的解空间和最优解:

(a) 外墙涂料的最大日需求量至多是 2.5 吨.

(b) 内墙涂料的日需求量至少是 2 吨.

(c) 内墙涂料的日需求量恰好比外墙涂料的日需求量多 1 吨.

(d) 原材料 M1 的日可用量至少是 24 吨.

(e) 原材料 M1 的日可用量至少是 24 吨, 并且内墙涂料的日需求量比外墙涂料的日需求量至少多 1 吨.

4. 某公司一天工作 10 个小时, 按 3 个连续的工艺生产 2 种产品. 下表列出了该问题数据:

产品	生产一个单位产品的分钟数			单位利润
	工艺 1	工艺 2	工艺 3	
1	10	6	8	2 美元
2	5	20	10	3 美元

确定这两种产品的最优生产策略.

***5.** 一家公司生产 A 和 B 两种产品. 产品 A 的销售量至少是产品 A 与产品 B 销售量总和的 80%. 然而, 公司每天不能销售多于 100 个单位的产品 A. 两种产品使用同一种原料, 并且最大日可用量是 240 磅. 生产每单位产品 A 需要用原料 2 磅, 生产每单位产品 B 需要用

原料 4 磅. 产品 A 和产品 B 的单位利润分别是 20 美元和 50 美元. 试为公司确定最优的生产策略.

6. Alumco 公司生产铝箔和铝棒. 预计最大生产能力是每天 800 张铝箔或者是每天 600 根铝棒. 最大日需求量是 550 张铝箔和 580 根铝棒. 利润为每张铝箔 40 美元, 每根铝棒 35 美元. 确定最优的日生产策略.

***7.** 某人打算把 5 000 美元投资在下一年度的两类项目: 投资项目 A 可盈利 5%, 投资项目 B 可盈利 8%. 市场研究表明, 至少应有 25% 的投资分配给项目 A, 至多应有 50% 的投资分配给项目 B. 进一步假定, 项目 A 的投资至少是项目 B 投资的一半. 将如何为这两类投资项目分配这笔专款?

8. Ozark 社区学院的继续教育系每学期共提供 30 门课程. 所提供的课程通常有两种类型: 一类是实用型的, 如木工手艺、文字处理和汽车维护; 另一类是人文型的, 如历史、音乐和美术. 为满足社区的需要, 每学期必须提供每种类型的课程至少 10 门. 继续教育系估计, 提供的实用和人文型课程的收入每门课分别大约是 1 500 美元和 1 000 美元.

(a) 为该学院设计最优课程设置.

(b) 证明, 每增加一门课程的价值是 1 500 美元, 相当于一门实用型课程的收入. 从增加课程的角度来看, 这一结果意味着什么?

9. ChemLabs 公司使用原料 I 和 II 生产两种家用清洁剂 A 和 B. 原料 I 和 II 的日可供应量分别是 150 单位和 145 单位. 一个单位的清洁剂 A 要消耗 0.5 个单位的原料 I 和 0.6 个单位的原料 II, 一个单位的清洁剂 B 要消耗 0.5 个单位的原料 I 和 0.4 个单位的原料 II. 每个单位的清洁剂 A 和 B 的利润分别是 8 美元和 10 美元. 清洁剂 A 的日需求量在 30 至 150 个单位之间, 清洁剂 B 的日需求量在 40 至 200 个单位之间. 试求清洁剂 A 和 B 的最优生产量.

10. 在 Ma&Pa 杂货商店, 货架的空间是有限的, 必须被有效地利用以提高利润. 有两种品牌的麦片, Grano 和 Wheatie, 共占据 60 平方英尺的货架空间. 一盒 Grano 麦片占 0.2 平方英尺, 一盒 Wheatie 麦片需要 0.4 平方英尺. Grano 和 Wheatie 麦片的最大日需求量分别是 200 盒和 120 盒. 卖一盒 Grano 净盈利 1 美元, 卖一盒 Wheatie 净盈利 1.35 美元. Ma&Pa 认为, 因为 Wheatie 的单位利润比 Grano 单位利润高 35%, 所以, 应该分配给 Wheatie 的空间比 Grano 的空间多 35%, 也就是将大约 57% 的空间分配给 Wheatie, 将大约 43% 的空间分配给 Grano. 你认为这样合理吗?

***11.** Jack 是 Ulern 大学的一名有抱负的新生. 他认识到劳逸结合的重要性, 因此, 他打算将每天可利用的 10 小时按一定的比例分配到工作和娱乐中. 他估计, 娱乐的快乐程度是工作的两倍. 他还打算, 学习的时间至少与娱乐的时间相同. 然而, Jack 认识到, 如果他打算完成留给他的全部作业, 他每天的娱乐时间不能超过 4 小时. Jack 该如何分配时间, 使得他从工作和娱乐中得到最大的快乐?

12. Wild West 生产两种类型的牛仔帽. 第 1 种类型帽子的劳动时间是第 2 种类型帽子的 2 倍. 如果将所有可利用的劳动时间全部用于生产第 2 种类型的帽子, 则公司每天可生产第 2 种类型帽子 400 顶. 两种类型帽子的市场限制分别是每天 150 顶和 200 顶. 第 1 种类型的帽子每顶盈利 8 美元, 第 2 种类型的帽子每顶盈利 5 美元. 确定每种类型的帽子生产多少能使利润达到最大.

13. Show & Sell 公司将它的产品在当地的广播电台和电视台做广告. 广告预算限制在每月 10 000 美元. 每分钟广播电台广告花费 15 美元, 每分钟电视广告耗资 300 美元. Show & Sell 希望做广播电台广告的时间至少是电视广告时间的 2 倍. 同时, 每月超过 400 分钟的广播广告是不现实的. 从过去的经验来看, 估计电视广告的效果是广播电台广告的 25 倍. 确定广播电台广告和电视广告的最优预算分配方案.

***14.** 怀俄明电力公司拥有蒸汽–涡轮发电厂. 因为怀俄明州储藏有丰富的煤, 公司用煤产生发电所需的蒸汽. 然而, 这可能导致排放物不符合环保署的标准. 环保署的法规限定, 每吨煤燃烧后二氧化硫的排放量不能超过 2 000 μL/L, 从发电厂排出的烟尘累计不能超过每小时 20 磅. 公司使用两种等级的煤粉 C1 和 C2 用于蒸汽发电. 这两种等级的煤粉在燃烧前通常被混合在一起. 为简单起见, 可以假定硫的排放量 (μL/L) 是两种煤粉混合比例的加权平均. 下列数据是基于两种等级的煤每小时 1 吨的消耗量所得出的.

煤的等级	硫的排放量 μL/L	烟尘的排放量 磅/小时	产生的蒸汽量 磅/小时
C1	1 800	2.1	12 000
C2	2 100	0.9	9 000

(a) 确定两种煤混合的最优比例.

(b) 求烟尘的排放量限制放松 1 磅对每小时产生蒸汽量的影响.

15. Top Toys 公司正计划在广播电台和电视台开展新的广告宣传活动. 做一次广播电台广告的花费是 300 美元, 一次电视广告的花费是 2 000 美元. 分配给这次活动的总预算是 20 000 美元. 然而, 为了保证在每种媒体上至少将有一次广播电台广告和一次电视广告, 每种媒体分配到的最大广告费不能超过总预算的 80%. 据估计, 第一次广播电台广告的听众将达到 5 000 人, 以后每增加一次广告, 新增听众只有 2 000 人. 对于电视广告, 第一次的广告观众将达到 4 500 人, 以后每增加一次广告, 新增观众 3 000 人. 如何将这笔预算在广播电台和电视台之间分配?

16. Burroughs 服装公司为 Walmark 廉价商店制造男式衬衫和女式上衣. Walmark 将接受由 Burroughs 提供的全部产品. 生产过程包括裁剪、缝纫和包装三部分. Burroughs 雇用 25 名工人在裁剪车间, 35 名工人在缝纫车间, 以及 5 名工人在包装车间. 工厂每天工作 8 小时, 每周工作 5 天. 下表给出了制造两种服装所需的时间及其利润:

服装	每件服装所需的时间 (分钟)			利润/件 (美元)
	裁剪	缝纫	包装	
男式衬衫	20	70	12	8
女式上衣	60	60	4	12

请为 Burroughs 公司确定每周最优的生产计划.

17. 某家具公司生产书桌和椅子. 切割车间为两种产品提供成品木材, 然后成品木材被送到组装车间组装. 组装好的家具再被送到油漆车间上漆. 切割车间每天能够生产 200 把椅子或 80 张书桌. 椅子组装车间每天可组装 120 把椅子, 书桌组装车间每天能组装 60 张书桌. 油漆车间每天能上漆 150 把椅子或者 110 张书桌. 已知每把椅子的利润是 50 美元, 每张书桌的利润是 100 美元. 为公司确定最优的生产组合.

*18. 一条装配线由 3 个相邻的工作站组成, 生产两种型号的收音机 HiFi-1 和 HiFi-2. 下表提供了 3 个工作站的装配时间.

工作站	每件产品所需的时间 (分钟)	
	HiFi-1	HiFi-2
1	6	4
2	5	5
3	4	6

对工作站 1, 2, 3 每天的维护所消耗的时间分别占每个站的总可利用时间的 10%, 14%, 12%, 每个工作站一天的最大可利用时间是 480 分钟. 确定最优的生产策略, 使 3 个工作站的总空闲 (或未使用) 时间最少.

19. TORA 实验. 将下面的线性规划输入到 TORA, 并选择图解法方式来展示线性规划图解屏幕.

$$
\begin{aligned}
\min\ z &= 3x_1+8x_2 \\
\text{s.t.}\ x_1 + x_2 &\geqslant 8 \\
2x_1 - 3x_2 &\leqslant 0 \\
x_1 + 2x_2 &\leqslant 30 \\
3x_1 - x_2 &\geqslant 0 \\
x_1 &\leqslant 10 \\
x_2 &\geqslant 9 \\
x_1, x_2 &\geqslant 0
\end{aligned}
$$

接下来, 在一张纸上, 画出所求问题的 x_1 和 x_2 轴, 并标上刻度 (也可以单击右边窗口上的 Print Graph, 得到一张标有刻度的纸). 现在, 在已准备好的纸上用手画出约束, 然后单击屏幕左边的窗口, 来检查你的答案. 对每个约束重复相同的操作, 最后画出目标函数, 终止整个程序. 通过 TORA 程序给出的立即反馈, 上述过程可用来检验和加强你对线性规划图解法的理解.

20. TORA 实验. 考虑下面的线性规划模型:

$$
\begin{aligned}
\max\ z &= 5x_1+4x_2 \\
\text{s.t.}\ 6x_1 + 4x_2 &\leqslant 24 \\
6x_1 + 3x_2 &\leqslant 22.5 \\
x_1 + x_2 &\leqslant 5 \\
x_1 + 2x_2 &\leqslant 6 \\
-x_1 + x_2 &\leqslant 1 \\
x_2 &\leqslant 2 \\
x_1, x_2 &\geqslant 0
\end{aligned}
$$

在线性规划中, 一个约束被称为是*冗余的*, 如果从模型中去掉该约束时, 可行解空间保持不变. 用 TORA 的作图功能找出冗余的约束, 然后说明去掉它们后 (或干脆不画它们) 不会影响解空间或最优解.

21. TORA 实验. 在 Reddy Mikks 模型中, 用 TORA 显示, 去掉原材料约束 [约束 (1) 和约束 (2)] 后将导致*无界解空间*. 在这种情况下, 模型的最优解会发生什么变化?

22. TORA 实验. 在 Reddy Mikks 模型中, 假定在问题中增加约束 $x_2 \geqslant 3$. 用 TORA 显示, 得到的模型有相互矛盾的约束, 这些约束不能同时被满足, 因此没有*可行解*.

2.2.2 极小化模型的解

例 2.2-2 (营养配方问题)

Ozark 农场每天至少使用 800 磅特殊饲料. 这种特殊饲料是由玉米和大豆粉配制而成, 并含有以下成分:

饲料	每磅饲料中含有成分的重量 (磅)		费用 (美元/磅)
	蛋白质	纤维	
玉　米	0.09	0.02	0.30
大豆粉	0.60	0.06	0.90

特殊饲料的营养要求是至少 30% 的蛋白质和至多 5% 的纤维. Ozark 农场希望确定每天最小成本的饲料配制.

因为饲料由玉米和大豆粉配制而成, 所以模型的决策变量定义为

$$x_1 = \text{每天混合饲料中玉米的重量 (磅)}$$

$$x_2 = \text{每天混合饲料中大豆粉的重量 (磅)}$$

目标函数是使得配制这种饲料的每天总成本 (美元) 最小, 因此表示为

$$\min\ z = 0.3x_1 + 0.9x_2$$

模型的约束反映饲料的日需要量和对营养成分的需求量. Ozark 农场一天至少需要饲料 800 磅, 相应的约束条件可以表示为

$$x_1 + x_2 \geqslant 800$$

对于蛋白质的营养需求约束, 含在 x_1 磅玉米和 x_2 磅大豆粉的蛋白质总量是 $(0.09x_1 + 0.6x_2)$ 磅. 这个量应至少等于总饲料混合 $(x_1 + x_2)$ 的 30%, 即

$$0.09x_1 + 0.6x_2 \geqslant 0.3(x_1 + x_2)$$

用类似的方法, 纤维的需求至多为 5%, 构造的约束是

$$0.02x_1 + 0.06x_2 \leqslant 0.05(x_1 + x_2)$$

化简约束, 将变量 x_1 和 x_2 移到不等式的左端, 只让常数保留在不等式的右端. 因此, 完整的模型变为

$$\begin{aligned} \min \quad & z = 0.3x_1 + 0.9x_2 \\ \text{s.t.} \quad & x_1 + \quad x_2 \geqslant 800 \\ & 0.21x_1 - 0.30x_2 \leqslant 0 \\ & 0.03x_1 - 0.01x_2 \geqslant 0 \\ & x_1,\ x_2 \geqslant 0 \end{aligned}$$

图 2.3 提供了模型图解的结果. 与 Reddy Mikks 模型 (例 2.2-1) 不同的是, 第 2 个和第 3 个约束通过原点. 为画出相应的直线, 我们需要另一个点, 这个点可以通过指定一个变量的值然后再解出另一个变量的值来获得. 例如, 在第 2 个约束中, $x_1 = 200$ 将产生 $0.21 \times 200 - 0.3x_2 = 0$ 或 $x_2 = 140$. 这意味着直线 $0.21x_1 - 0.3x_2 = 0$ 通过点 (0,0) 和点 (200,140). 还需要注意的一点是, (0,0) 点不能用作第 2 个约束和第 3 个约束的参考点, 因为这两条直线均通过原点. 不过, 我们可以用任何其他点 [例如, (100,0) 或 (0,100)] 作为参考点.

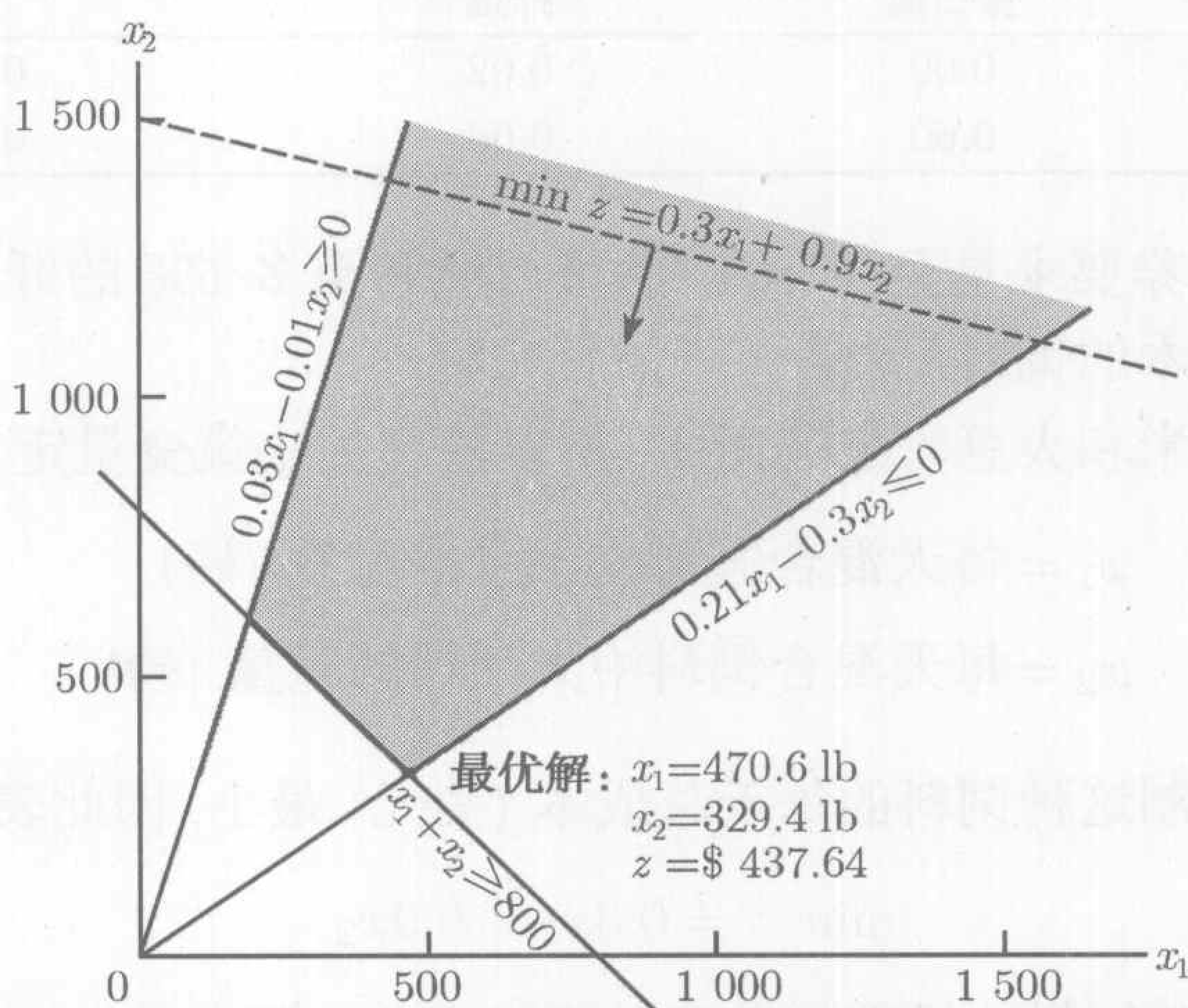


图 2.3　营养配方模型的图解

解　因为现在的模型寻找目标函数的极小值, 我们需要尽可能地沿图 2.3 所示的方向减少 z 值. 最优解位于两条直线 $x_1 + x_2 = 800$ 和 $0.21x_1 - 0.3x_2 = 0$ 的交点处, 得到 $x_1 = 470.59$ 磅和 $x_2 = 329.41$ 磅. 配制这种饲料的最小成本为 $z = 0.3 \times 470.59 + 0.9 \times 329.42 = 437.65$ 美元/天.

评注　需要注意的是问题约束的构造方法. 因为模型是极小化总成本, 我们可能会认为, 寻求的解应该恰好是 800 磅饲料. 的确, 这正是上述最优解给出的结果. 这是否意味着将第一个约束去掉, 而简单地用总量 800 磅代替, 并保留其他的约束呢? 为得到答案, 我们把新的蛋白质和纤维约束表示成

$$0.09x_1 + 0.6x_2 \geqslant 0.3 \times 800$$

$$0.02x_1+0.06x_2\leqslant 0.05\times 800$$

即

$$0.09x_1+0.6x_2\geqslant 240$$

$$0.02x_1+0.06x_2\leqslant 40$$

这组新的约束产生的解是 $x_1=0$ 和 $x_2=400$ 磅 (通过 TORA 验证!), 它并不满足隐含的 800 磅饲料的需求. 这意味着必须明确采用约束 $x_1+x_2\geqslant 800$, 并且蛋白质和纤维的约束必须严格地保持原来的表示方法.

按照前面的思维方法, 可以试着用约束 $x_1+x_2=800$ 代替约束 $x_1+x_2\geqslant 800$. 对于前面的例子, 两种约束产生相同的结果. 但推广到一般问题, 情况并非如此. 例如, 假设每天的混合饲料至少包含 500 磅玉米. 在这种情况下, 最优解是要求使用 500 磅的玉米和 350 磅的大豆粉 (通过 TORA 验证!), 它等价于每天的饲料混合量是 $500+350=850$. 强迫使用等式约束将导致问题没有可行解 (通过 TORA 验证!). 另一方面, 不等式同时也包含了相等的情况, 因此, 它并没有排除模型产生恰好 800 磅的饲料混合, 所以应该保留不等式约束. 上述讨论说明, 我们不应该"预先猜测"其结果而强迫附加等式限制, 我们总是应该使用不等式约束, 除非是明确规定使用等式约束的情形.

习题 2.2B

1. 确定下列情况中 z 的下降方向:

*(a) min $z=4x_1-2x_2$ (b) min $z=-3x_1+x_2$ (c) min $z=-x_1-2x_2$

2. 对于营养配方模型, 假定玉米的日可利用量不超过 450 磅. 找出新的解空间, 并确定新的最优解.

3. 对于营养配方模型, 如果饲料混合不超过一天 800 磅, 模型将得到什么样的最优解? 这样的解有意义吗?

4. John 在上学期间, 每周要工作 20 小时来勤工俭学. 他在两个零售店里找到了工作. 在第 1 个零售店, 他可以一周工作 5 至 12 个小时; 第 2 个零售店允许他工作 6 至 10 小时. 两个店支付相同的小时工资. 考虑要在每个店工作多少时间时, John 打算依据每个店的工作强度来作出决定. 通过与现在的店员交谈, 并用 1 至 10 之间的数来度量递增的工作强度, John 估计第 1 个店和第 2 个店的强度分别是 8 和 6. 因为强度是随工作时间增加的, 他假定, 每个店在一周结束时的总强度, 与他在该店工作的小时数成正比. 那么 John 应该在每个店工作多少小时呢?

***5.** OilCo 公司正在建设一个生产 4 种产品 (柴油、汽油、润滑油和喷气式飞机燃料) 的精炼厂. 这些产品每种的最小需求量 (桶/天) 分别是 14 000, 30 000, 10 000 和 8 000. 伊朗和阿联酋按照合同通过船只为 OilCo 提供原油. 由于原油受 OPEC(Organization of Petroleum Exporting Countries, 石油输出国组织) 配额的限制, 新的精炼厂至少有 40% 的原油来自

伊朗, 其余来自阿联酋. OilCo 公司预测, 石油产品的需求和原油配额将在今后的 10 年中保持稳定.

两种原油规格导致不同的产品系列：一桶伊朗原油可生产出 0.2 桶柴油、0.25 桶汽油、0.1 桶润滑油和 0.15 桶喷气式飞机燃油; 阿联酋原油的相应指标分别是 0.1, 0.6, 0.15, 0.1. OilCo 需要确定精炼厂的最小生产能力 (桶/天).

6. 某股民打算投入一批资金, 用它产生至少 10 000 美元的年收益. 可以用这笔资金购买两种股票：蓝筹股和高科技股. 这两种股票的平均年收益分别是 10% 和 25%. 虽然高科技股的年收益较高, 但它风险更大, 因此他打算限制在这种股票的投资不超过总投资额的 60%. 他对每种股票的最少投资量应该是多少才能完成他的投资目标?

***7.** 一家工业再生中心用 A 和 B 两种废铝金属生产特殊合金. 废料 A 含有 6% 的铝、3% 的硅和 4% 的碳. 废料 B 含有 3% 的铝、6% 的硅和 3% 的碳. 每吨的废料 A 和 B 的成本分别是 100 美元和 80 美元. 这种特殊合金的规格要求是：(1) 铝的含量必须在 3% 至 6% 之间; (2) 硅的含量必须在 3% 至 5% 之间; (3) 碳的含量必须在 3% 至 7% 之间. 求用于生产 1 000 吨这种特殊合金所需废料的最优混合.

8. TORA 实验. 考虑营养配方模型, 令目标函数如下：

$$\min \ z = 0.8x_1 + 0.8x_2$$

用 TORA 软件演示最优解在*两个*不同的角点处取得, 并且两个点有相同的目标值. 在这种情况下, 称问题有*可选择最优解*. 解释导致这种情形的条件, 并说明, 该问题实际上有无穷多个可选择的最优解, 然后给出求所有这些最优解的公式.

2.3 线性规划应用选讲

本节介绍若干实用的线性规划模型, 在这些模型中, 变量的定义、目标函数和约束的构造并不像二维变量模型那么简单. 这些应用问题所涉及的领域包括有：(1) 城市规划; (2) 货币套汇; (3) 投资; (4) 生产计划和库存控制; (5) 石油混合与提炼; (6) 人力计划.

本节详细讨论每一种模型, 并对模型的最优解进行分析并给出解释.

2.3.1 城市规划[①]

城市规划要解决 3 大问题：(1) 建造新的住宅开发; (2) 改造城市内部的老旧房屋和娱乐场所; (3) 规划公共设施 (如学校和机场). 与这些项目相关的约束包括经济 (土地、建筑物、资金) 与社会 (学校、停车场、收入水平) 两个方面. 在城市规划中目标并不是固定的, 例如, 在新的住宅开发项目中, 利润通常是承担项目的动机; 在其他的两类项目中, 目标包括了社会、政治、经济和文化诸多需要考虑的因素. 又如, 在 2004 年引起公众所注意的事件中, 俄亥俄州的一名市长想要征用一片旧城区, 用来开发豪华住宅. 其动机是增加税收以帮助弥补经费预算的不足. 本节介绍的例子就是按照俄亥俄州的案例编写的.

① 本节基于文献 Laidlaw(1972).

例 2.3-1 (城市更新模型)

Erstville 市面临着严重的预算不足问题. 为寻求一种长期的解决方案, 市议会投票决定, 征用一块内城的住宅区域, 进行一项现代化的开发项目, 以增加税收来源.

改造工程包括两个阶段: (一) 拆除不符合标准的住宅, 为新的开发提供土地; (二) 建造新的建筑. 下面是情况概要.

(1) 拆除大约 300 套不符合标准的住宅. 每套住宅占地 0.25 英亩. 拆除一套征地住宅的成本是 2 000 美元.

(2) 新的单、双、三和四户住宅 (单元) 的土地面积分别是 0.18、0.28、0.4 和 0.5 英亩. 街道、开阔地和公共设施占可利用面积总量的 15%.

(3) 在新的开发项目中, 三户与四户的住宅单元数的总和至少占总住宅单元的 25%. 单户住宅单元数至少应占总单元数的 20%. 双户住宅单元数至少占总单元数的 10%.

(4) 对于单、双、三和四户住宅, 每单元征税额分别是 1 000 美元、1 900 美元、2 700 美元和 3 400 美元.

(5) 对于单、双、三和四户住宅, 每单元的建筑成本分别是 50 000 美元、70 000 美元、130 000 美元和 160 000 美元. 通过当地银行筹措资金总计最高达到 1 500 万美元.

应建多少单元各种类型的住宅使得税收总额达到最大?

数学模型 除了确定建造每种类型住宅单元的数量外, 我们还需要确定有多少旧房屋必须拆除, 为新的开发项目提供场地. 因此, 问题的变量可定义如下:

$$x_1 = \text{建造单户住宅的单元数}$$

$$x_2 = \text{建造双户住宅的单元数}$$

$$x_3 = \text{建造三户住宅的单元数}$$

$$x_4 = \text{建造四户住宅的单元数}$$

$$x_5 = \text{拆除旧住宅的单元数}$$

目标函数是从所有 4 种类型的住宅中使得总的税收达到最大, 即

$$\max\ z = 1\,000x_1 + 1\,900x_2 + 2\,700x_3 + 3\,400x_4$$

问题的第一个约束是土地的可用量:

$$\text{用于新建住宅面积} \leqslant \text{净可用面积}$$

从问题的数据我们得到

$$\text{新建住宅需要的面积} = 0.18x_1 + 0.28x_2 + 0.4x_3 + 0.5x_4$$

为确定可利用面积, 被拆除的旧住宅每套占 0.25 英亩土地, 因此, 得到 $0.25x_5$ 英亩. 允许用总量的 15% 用于开阔区域、街道和活动设施, 净可利用面积是 $0.85(0.25x_5) = 0.2125x_5$. 得到的约束是

$$0.18x_1 + 0.28x_2 + 0.4x_3 + 0.5x_4 \leqslant 0.212\,5x_5$$

即

$$0.18x_1 + 0.28x_2 + 0.4x_3 + 0.5x_4 - 0.212\,5x_5 \leqslant 0$$

被拆除住宅的数量不能超过 300 套, 将它写成

$$x_5 \leqslant 300$$

下面加上各种类型住宅单元数量的限制约束:

单户住宅单元的数量 ⩾ 总单元数量的 20%
双户住宅单元的数量 ⩾ 总单元数量的 10%
三户与四户住宅单元数量之和 ⩾ 总单元数量的 25%

这些约束转述成数学表达式就是

$$\begin{aligned} x_1 &\geqslant 0.2(x_1 + x_2 + x_3 + x_4) \\ x_2 &\geqslant 0.1(x_1 + x_2 + x_3 + x_4) \\ x_3 + x_4 &\geqslant 0.25(x_1 + x_2 + x_3 + x_4) \end{aligned}$$

还剩下一个约束就是要保证拆除和建造的费用在允许的预算之内, 即

拆除和建造费用的总和 ⩽ 可用预算

以 1 000 美元表示费用的单位, 得到

$$(50x_1 + 70x_2 + 130x_3 + 160x_4) + 2x_5 \leqslant 15\,000$$

因此, 完整的模型如下:

$$\begin{aligned} \max\ z = 1\,000x_1 + 1\,900x_2 + 2\,700x_3 + 3\,400x_4 \\ \text{s.t. } 0.18x_1 + 0.28x_2 + 0.4x_3 + 0.5x_4 - 0.212\,5x_5 &\leqslant 0 \\ x_5 &\leqslant 300 \\ -0.8x_1 + 0.2x_2 + 0.2x_3 + 0.2x_4 &\leqslant 0 \\ 0.1x_1 - 0.9x_2 + 0.1x_3 + 0.1x_4 &\leqslant 0 \\ 0.25x_1 + 0.25x_2 - 0.75x_3 - 0.75x_4 &\leqslant 0 \\ 50x_1 + 70x_2 + 130x_3 + 160x_4 + 2x_5 &\leqslant 15\,000 \\ x_1, x_2, x_3, x_4, x_5 &\geqslant 0 \end{aligned}$$

解 最优解 (用文件 amplEX2.3-1.txt 或 solverEx2.3-1.xls 计算得到) 为

$$\text{总税收额} = z = 343\,965(\text{美元})$$

$$单户住宅数量 = x_1 = 35.83 \approx 36\ (套)$$
$$双户住宅数量 = x_2 = 98.53 \approx 99\ (套)$$
$$三户住宅数量 = x_3 = 44.79 \approx 45\ (套)$$
$$四户住宅数量 = x_4 = 0\ (套)$$
$$拆除旧住宅数量 = x_5 = 244.49 \approx 245\ (套)$$

评注 线性规划并不能自动地保证得到整数解, 这就是将连续的值经过四舍五入得到整数的原因. 这个四舍五入的结果要求建造 $180 (= 36 + 99 + 45)$ 套单元, 拆除 245 套旧住宅, 税收总额是 345 600 美元. 但需要记住的是, 在一般情况下, 由四舍五入得到的结果不一定是可行的. 事实上, 现在的四舍五入的结果就破坏了 70 000 美元的预算约束 (请验证!). 有趣的是, 真实的最优整数解 (用第 8 章的算法) 是 $x_1 = 36$, $x_2 = 98$, $x_3 = 45$, $x_4 = 0$ 和 $x_5 = 245$, 并且最优值 $z = 343\,700$ 美元. 细心的读者可注意到, 四舍五入的结果产生更好的目标最优值, 这似乎是矛盾的. 但究其原因是, 四舍五入的结果要多建一套双户住宅单元, 这只有在预算超过 70 000 美元才是可行的.

习题 2.3A

1. 某房地产经纪人正计划开发出租住宅和零售商业区的业务. 住宅区由简易公寓、联式和单户住宅组成. 估计潜在承租人的最大需求是 500 套简易公寓、300 套联式住宅和 250 套单户住宅, 但联式住宅的数量必须至少等于简易公寓和单户住宅数量的 50%. 零售店面积与住宅的套数成比例, 对于简易公寓、联式住宅和单户住宅, 其比例分别至少是 10 平方英尺、15 平方英尺和 18 平方英尺. 然而, 由于土地可用性限制零售店面积不能超过 10 000 平方英尺. 对于简易公寓、联式住宅的单户住宅, 估计月租收入分别是 600 美元、750 美元和 1 200 美元. 零售店面的租金是 100 美元/平方英尺. 确定最优零售店面积和居民住户的数量.

2. Fayetteville 市议会正在审批建设新的 200 000 平方英尺会议中心的议案. 已提出了两处地点, 均需要行使"征用权"的法律才能获得地产. 下表给出了两个地点 (及附近) 的有关地产, 以及购置成本.

地产	地点 1		地点 2	
	面积 (1 000 英尺^2)	成本 (1 000 美元)	面积 (1 000 英尺^2)	成本 (1 000 美元)
1	20	1 000	80	2 800
2	50	2 100	60	1 900
3	50	2 350	50	2 800
4	30	1 850	70	2 500
5	60	2 950		

允许获得部分产权. 如果选择地点 1, 至少要获得产权 4 的 75%; 若选择地点 2, 至少要获得产权 3 的 50%. 虽然地点 1 的单位面积产权要贵些, 但建筑成本低于地点 2, 因为地点

1 的基础设施更好. 地点 1 的建筑成本是 2 500 万美元, 地点 2 的建筑成本是 2 700 万美元. 应该选择哪一个地点, 获得哪些地产?

***3.** 某城市在未来的 5 年内将启动 4 个城市住房改造工程. 每项工程有不同的开始时间, 工程周期也不一样. 下表提供这些项目的基本数据 (单位: 100 万美元):

	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年	总费用	年收入
工程 1	开始		结束			5.0	0.05
工程 2		开始			结束	8.0	0.07
工程 3	开始				结束	15.0	0.15
工程 4			开始	结束		1.2	0.02
预算	3.0	6.0	7.0	7.0	7.0		

工程 1 和工程 4 必须在规定的周期内全部完成. 必要时, 其余的两项工程可以在预算的限制内完成一部分. 然而, 每项工程在它的规定时间内必须至少完成 25%. 每年年底, 工程完成的部分立刻入住, 并且实现一定比例的收入. 例如, 如果工程 1 在第 1 年完成 40%, 在第 3 年完成剩下的 60%, 在 5 年计划范围内的相应收入是 $0.4 \times \$50\ 000$(第 2 年)$+0.4 \times \$50\ 000$(第 3 年)$+(0.4+0.6) \times \$50\ 000$ (第 4 年)$+(0.4+0.6) \times \$50\ 000$(第 5 年)$= (4 \times 0.4 + 2 \times 0.6) \times \$50\ 000$. 试为工程确定最优的时间进度表, 使得 5 年内的总收入达到最大. 为简便起见, 不考虑货币的时间价值.

4. Fayetteville 市正计划着手一项城市改造工程, 其工程包括为中低收入居民建造的排式住宅, 为高收入者建造的豪华公寓, 以及公共建筑. 也包括公立小学和商业零售设施. 学校的大小 (教室的间数) 与学生的数量成比例, 而且零售设施的建筑面积与住宅单元的数量成正比. 下表提供了该项工程的相关数据:

	低收入	中等收入	高收入	公共住宅	学校教室	零售设施
最小单元数	100	125	75	300		0
最大单元数	200	190	260	600		25
每单元土地面积 (英亩)	0.05	0.07	0.03	0.025	0.045	0.1
每单元的学生平均数	1.3	1.2	0.5	1.4		
每单元零售设施需求 (英亩)	0.023	0.034	0.046	0.023	0.034	
每单元的年收入 (美元)	7 000	12 000	20 000	5 000	—	15 000

按照每间教室至多 25 名学生, 新学校最大可占地 2 英亩. 学校每间教室年运行费用是 10 000 美元. 工程将建于一片 50 英亩的空地上, 其产权归市政府所有. 另外, 该工程能够征用邻近的、被划定为贫民区的 200 套住宅的产权. 每套贫民住宅占地 0.25 英亩. 购买和拆除一套贫民住宅的费用是 7 000 美元. 空地、街道和停车场用去总可利用土地的 15%. 试设计一个线性规划, 确定该项工程的最优计划.

5. Realco 公司拥有 800 英亩未开发土地, 坐落在 Ozark 山中心区景色优美的湖边. 过去, 对于湖周边的新开发简直是没有规划. 现在, 湖岸星星点点地布满了度假住宅和化粪池, 绝大多数化粪池建造得不合理, 但普遍都在使用. 多年来, 来自化粪池的渗流导致了严重的水污染. 为了抑制湖水的进一步恶化, 县级官员着眼于未来的发展颁布了严厉的法令: (1) 只能建造单户、双户和三户住宅, 并且单户住宅至少占总住宅的 50%; (2) 为了要限制化粪池的

数量, 单户、双户和三户住宅所占面积最少应分别是 2 英亩、3 英亩和 4 英亩; (3) 娱乐区的建造比例是, 每 200 个家庭要建设 1 英亩的娱乐区; (4) 为了保护该湖的生态环境, 不能抽取地下水用于住宅或花园. Realco 的董事长正在研究开发 800 英亩房地产项目的方案. 这一新开发项目将包括单户、双户和三户住宅. 估计有 15% 的土地将会用于街道和公用设施的附属建筑物. Realco 估计不同住宅单元的收益如下：

住宅单元	单户	双户	三户
每单元净收益 (美元)	10 000	12 000	15 000

连接住宅区域的自来水系统的费用与建造的单元数量成比例. 然而, 县政府对于工程的最低收费是 100 000 美元. 另外, 对现有的供水系统的扩建规模, 每天高峰时段新增的供水量不超过 200 000 加仑. 下面的数据列出了自来水接入的费用, 以及一般家庭的用水量：

住宅单元	单户	双户	三户	娱乐
每单元自来水的接入费用 (美元)	1 000	1 200	1 400	800
每单元用水量 (加仑/天)	400	600	840	450

为 Realco 制定一项最优规划.

6. 考虑习题 5 的 Realco 模型. 假定还可以购买到另一块 100 英亩的土地, 其费用是 450 000 美元, 这样总面积增加到 900 英亩. 这样做对 Realco 有利可图吗?

2.3.2 套汇①

在当今的全球经济中, 跨国公司必须要处理其业务所在国的货币. 套汇, 指在不同市场中同时进行货币的购买和销售, 使得资金由一种货币在转换成另一种货币的运作中具有获利机会. 例如, 在 2001 年要将 1 000 英镑兑换成美元, 按 1.6 美元兑换 1 英镑的兑换率, 能换到 1 600 美元. 另一种兑换的方法是先将英镑兑换成日元, 再将日元兑换成美元, 在 2001 年的兑换率是 1 英镑 =175 日元, 1 美元 =106 日元. 这样, 能得到的美元数量是 $\frac{(1\ 000\text{英镑}\times 175\text{日元})}{105\text{日元}} = 1\ 666.67$ 美元. 这个例子显示了先将英镑兑换成日元, 再将日元兑换成美元的好处. 这一节将介绍如何将涉及多种货币的套汇问题设计成一个线性规划, 以及如何求解.

例 2.3-2 (套汇模型)

假定一家公司总共有 500 万美元, 可以兑换成欧元 (€)、英镑 (£)、日元 (¥) 和科威特第纳尔 (KD). 货币经销商规定下列单笔货币交易量的上限：500 万美元, 300 万欧元, 350 万英镑, 1 亿日元和 280 万科威特第纳尔. 下表提供典型的现汇率. 对角线以下的比率是对角线以上比率的倒数. 例如, 汇率 (€ → \$) = 1/汇率 (\$ → €) = 1/0.769 = 1.30.

① 本节基于文献 J. Kornbluth and G. Salkin (1987) 的第 6 章.

	\$	€	£	¥	KD
\$	1	0.769	0.625	105	0.342
€	$\frac{1}{0.769}$	1	0.813	137	0.445
£	$\frac{1}{0.625}$	$\frac{1}{0.813}$	1	169	0.543
¥	$\frac{1}{105}$	$\frac{1}{137}$	$\frac{1}{169}$	1	0.003 2
KD	$\frac{1}{0.342}$	$\frac{1}{0.445}$	$\frac{1}{0.543}$	$\frac{1}{0.003\ 2}$	1

通过货币市场的循环, 增加美元的持有量 (初始值为 500 万美元) 是可能的吗?

数学模型 以 500 万美元的情形开始. 这笔美元, 在最终又被兑换成美元之前, 经历一系列兑换成其他货币的过程. 因此, 问题是试图确定每种货币的兑换量, 使总的美元持有数量达到最大.

为了建立这个模型以及简化符号, 用下列数字编码表示各种货币.

货币	\$	€	£	¥	KD
代码	1	2	3	4	5

定义

$$x_{ij} = \text{货币 } i \text{ 兑换成货币 } j \text{ 的数量}, i, j = 1, 2, \cdots, 5$$

例如, x_{12} 是美元兑换成欧元的数量, x_{51} 是科威特第纳尔兑换成美元的数量. 我们进一步定义另外两个变量, 表示套汇问题的输入与输出:

$$I = \text{初始美元数量} \ (= 500 \text{ 万美元})$$

$$y = \text{最终持有的美元数量} \ (\text{由解来确定})$$

我们的目标是求最终持有的美元数量 y 的最大值, 所受的约束是货币流通的限制和不同交易的最高限额.

首先建立模型的约束条件, 图 2.4 显示了将美元兑换成英镑的情形. 以货币 1 开始的美元数量为 x_{13}, 兑换成英镑数量是 $0.625x_{13}$, 即以货币 3 结束. 同时, 美元的交易量不能超过经销商设定的限额, 即 $x_{13} \leqslant 5$.

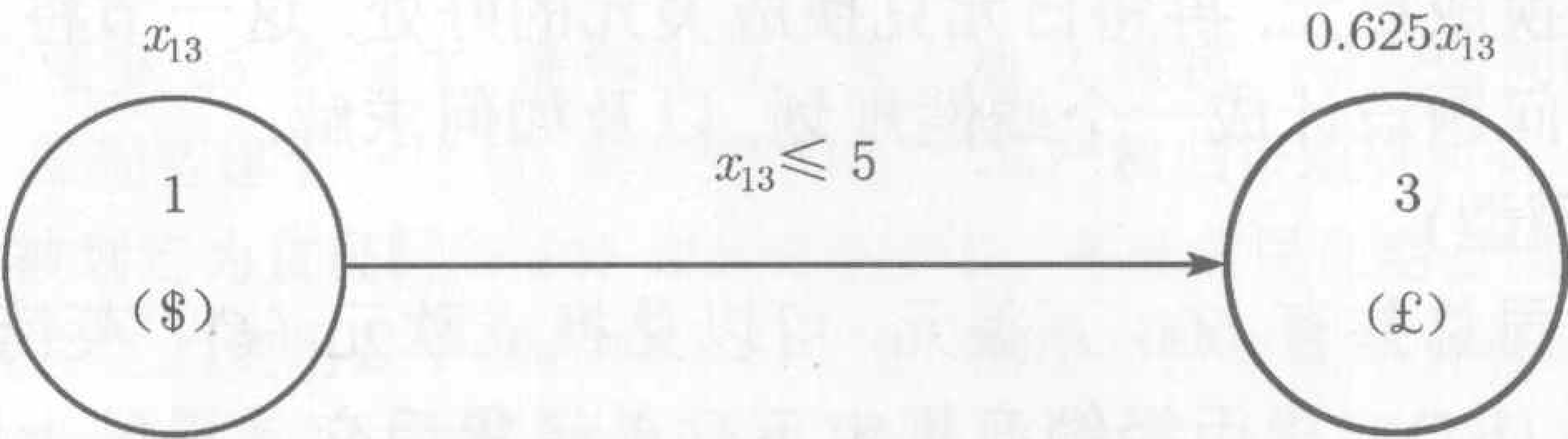


图 2.4 美元与英镑之间输入/输出变量 x_{13} 的定义

为了使资金流从一种货币兑换到另一种货币时守恒, 每种货币必须满足下列输入–输出方程:

$$\text{可用货币总金额 (输入)} = \text{兑换成其他货币的总金额 (输出)}$$

(1) 美元($i=1$):

$$\begin{aligned}\text{可用美元总数} &= \text{初始美元数量} + \text{从其他货币兑换得到的美元数量}\\ &= I + (€\to \$) + (£\to \$) + (¥\to \$) + (\text{KD}\to \$)\\ &= I + \frac{1}{0.769}x_{21} + \frac{1}{0.625}x_{31} + \frac{1}{105}x_{41} + \frac{1}{0.342}x_{51}\end{aligned}$$

$$\begin{aligned}\text{美元分布总数} &= \text{最终持有的美元数量} + \text{美元兑换成其他货币的数量}\\ &= y + (\$\to €) + (\$\to £) + (\$\to ¥) + (\$\to \text{KD})\\ &= y + x_{12} + x_{13} + x_{14} + x_{15}\end{aligned}$$

已知 $I=5$, 因此, 美元约束变为

$$y + x_{12} + x_{13} + x_{14} + x_{15} - \left(\frac{1}{0.769}x_{21} + \frac{1}{0.625}x_{31} + \frac{1}{105}x_{41} + \frac{1}{0.342}x_{51}\right) = 5$$

(2) 欧元($i=2$):

$$\begin{aligned}\text{可用欧元总数} &= (\$\to €) + (£\to €) + (¥\to €) + (\text{KD}\to €)\\ &= 0.769x_{12} + \frac{1}{0.813}x_{32} + \frac{1}{137}x_{42} + \frac{1}{0.445}x_{52}\end{aligned}$$

$$\begin{aligned}\text{欧元分布总数} &= (€\to \$) + (€\to £) + (€\to ¥) + (€\to \text{KD})\\ &= x_{21} + x_{23} + x_{24} + x_{25}\end{aligned}$$

因此, 欧元约束是

$$x_{21} + x_{23} + x_{24} + x_{25} - \left(0.769x_{12} + \frac{1}{0.813}x_{32} + \frac{1}{137}x_{42} + \frac{1}{0.445}x_{52}\right) = 0$$

(3) 英镑($i=3$):

$$\begin{aligned}\text{可用英镑总数} &= (\$\to £) + (€\to £) + (¥\to £) + (\text{KD}\to £)\\ &= 0.625x_{13} + 0.813x_{23} + \frac{1}{169}x_{43} + \frac{1}{0.543}x_{53}\end{aligned}$$

$$\begin{aligned}\text{英镑分布总数} &= (£\to \$) + (£\to €) + (£\to ¥) + (£\to \text{KD})\\ &= x_{31} + x_{32} + x_{34} + x_{35}\end{aligned}$$

因此, 英镑约束是

$$x_{31} + x_{32} + x_{34} + x_{35} - \left(0.625x_{13} + 0.813x_{23} + \frac{1}{169}x_{43} + \frac{1}{0.543}x_{53}\right) = 0$$

(4) 日元($i=4$):

$$\begin{aligned}\text{可用日元总数} &= (\$\to ¥) + (€\to ¥) + (£\to ¥) + (\text{KD}\to ¥)\\ &= 105x_{14} + 137x_{24} + 169x_{34} + \frac{1}{0.003\,2}x_{54}\end{aligned}$$

$$\begin{aligned}\text{日元分布总数} &= (¥\to \$) + (¥\to €) + (¥\to £) + (¥\to \text{KD})\\ &= x_{41} + x_{42} + x_{43} + x_{45}\end{aligned}$$

因此, 日元约束是

$$x_{41}+x_{42}+x_{43}+x_{45}-(105x_{14}+137x_{24}+169x_{34}+\frac{1}{0.003\ 2}x_{54})=0$$

(5) 科威特第纳尔 ($i=5$):

$$\begin{aligned}\text{可用 KD 总数} &= (\$ \to \text{KD}) + (€ \to \text{KD}) + (£ \to \text{KD}) + (¥ \to \text{KD}) \\ &= 0.342x_{15}+0.445x_{25}+0.543x_{35}+0.003\ 2x_{45} \\ \text{KD 分布总数} &= (\text{KD} \to \$) + (\text{KD} \to €) + (\text{KD} \to £) + (\text{KD} \to ¥) \\ &= x_{51}+x_{52}+x_{53}+x_{54}\end{aligned}$$

因此, KD 约束是

$$x_{51}+x_{52}+x_{53}+x_{54}-(0.342x_{15}+0.445x_{25}+0.543x_{35}+0.003\ 2x_{45})=0$$

还剩下的约束是最高交易限额, 即 500 万美元, 300 万欧元, 350 万英镑, 1 亿日元和 280 万 KD. 这些可以表达成

$$\begin{aligned} &x_{1j}\leqslant 5,\quad j=2,3,4,5 \\ &x_{2j}\leqslant 3,\quad j=1,3,4,5 \\ &x_{3j}\leqslant 3.5,\quad j=1,2,4,5 \\ &x_{4j}\leqslant 100,\quad j=1,2,3,5 \\ &x_{5j}\leqslant 2.8,\quad j=1,2,3,4\end{aligned}$$

现在给出完整的模型:

$$\begin{aligned} \max\ & z=y \\ \text{s.t.}\ & y+x_{12}+x_{13}+x_{14}+x_{15}-(\frac{1}{0.769}x_{21}+\frac{1}{0.625}x_{31}+\frac{1}{105}x_{41}+\frac{1}{0.342}x_{51})=5 \\ & x_{21}+x_{23}+x_{24}+x_{25}-(0.769x_{12}+\frac{1}{0.813}x_{32}+\frac{1}{137}x_{42}+\frac{1}{0.445}x_{52})=0 \\ & x_{31}+x_{32}+x_{34}+x_{35}-(0.625x_{13}+0.813x_{23}+\frac{1}{169}x_{43}+\frac{1}{0.543}x_{53})=0 \\ & x_{41}+x_{42}+x_{43}+x_{45}-(105x_{14}+137x_{24}+169x_{34}+\frac{1}{0.003\ 2}x_{54})=0 \\ & x_{51}+x_{52}+x_{53}+x_{54}-(0.342x_{15}+0.445x_{25}+0.543x_{35}+0.003\ 2x_{45})=0 \\ & x_{1j}\leqslant 5,\quad j=2,3,4,5 \\ & x_{2j}\leqslant 3,\quad j=1,3,4,5 \\ & x_{3j}\leqslant 3.5,\quad j=1,2,4,5 \\ & x_{4j}\leqslant 100,\quad j=1,2,3,5 \\ & x_{5j}\leqslant 2.8,\quad j=1,2,3,4 \\ & x_{ij}\geqslant 0,\quad \text{对于所有的 } i \text{ 和 } j \end{aligned}$$

解 最优解 (用文件 amplEx2.3-2.txt 或 solverEx2.3-2.xls) 为

解	含义解释
$y=5.090\,32$	最终持有美元数量 = 5 090 320 美元, 净美元利润 = 90 320 美元, 收益率为 1.8064%
$x_{12}=1.462\,06$	购买价值 1 462 060 美元的欧元
$x_{15}=5$	购买价值 5 000 000 美元的科威特第纳尔
$x_{25}=3$	购买价值 3 000 000 欧元的科威特第纳尔
$x_{31}=3.5$	购买价值 3 500 000 英镑的美元
$x_{32}=0.931\,495$	购买价值 931 495 英镑的欧元
$x_{41}=100$	购买价值 100 000 000 日元的美元
$x_{42}=100$	购买价值 100 000 000 日元的欧元
$x_{43}=100$	购买价值 100 000 000 日元的英镑
$x_{53}=2.085$	购买价值 2 085 000 KD 的英镑
$x_{54}=0.96$	购买价值 960 000 KD 的日元

评注 起初, 可能感觉这个结果是荒谬的, 因为它要求用 $x_{12}+x_{15}=1.462\,06+5=6.462\,06$, 即用 6 462 060 美元去购买欧元和科威特第纳尔, 但初始美元数量只有 5 000 000 美元. 额外的美元是从哪儿来的呢? 在实践中, 所发生的正是, 将已知结果按某种次序提交给货币经销商, 这意味着在购买之前, 我们并不需要等待直到某种货币积累到足够量的程度. 最后, 所有这些交易的净结果是将 5 000 000 美元的净成本交给投资者. 将结果中所有的美元交易加在一起即可得到:

$$
\begin{aligned}
I &= y+x_{12}+x_{13}+x_{14}+x_{15}-\left(\tfrac{1}{0.769}x_{21}+\tfrac{1}{0.625}x_{31}+\tfrac{1}{105}x_{41}+\tfrac{1}{0.342}x_{51}\right)\\
&=5.090\,32+1.462\,06+5-\left(\tfrac{3.5}{0.625}+\tfrac{100}{105}\right)\\
&=5
\end{aligned}
$$

注意, x_{21}, x_{31}, x_{41} 和 x_{51} 分别是欧元、英镑、日元和科威特第纳尔的数量, 因此必须兑换成为美元.

习题 2.3B

1. 修改套汇模型, 要求佣金等于任何货币购买量的 0.1%. 假定佣金并不影响正在循环的资金, 在完成了整个的兑换后, 再交纳佣金. 与原来的模型相比, 其结果有何不同?

***2.** 假定该公司愿意将 500 万美元兑换成能够提供最高收益率的任何货币. 修改原始模型, 确定哪种货币最好.

3. 假定初始的货币量 $I=700$ 万美元, 公司打算最优地将它兑换成欧元、英镑和日元的组合. 最终的混合中最多有 200 万欧元、300 万英镑和 2 亿日元. 修改原始模型, 确定这 3 种货币的最优购买组合.

4. 假定公司希望购买 600 万美元. 不同种货币的交易上限与原问题相同. 为这次交易设计一个购买清单, 已知现有的混合资金不超过 300 万欧元、200 万英镑和 200 万科威特第纳尔.

5. 假定公司有 200 万美元、500 万欧元、400 万英镑. 设计一个买–卖次序, 提高全部兑换成日元的货币总持有量.

2.3.3 投资

今天的投资者面临着多种投资机会. 投资问题的例子包括工程的资金预算、债券投资策略、股票投资组合的选择, 以及银行贷款政策的制定. 在许多这些问题中, 都可以利用线性规划来选择最优的投资机会组合, 在满足投资商设定的投资条件的同时, 使得收益达到最大.

例 2.3-3 (贷款策略模型)

Thriftem 银行正在制定一项总额可达 1 200 万美元的贷款策略. 下表提供了各类贷款的相关数据.

贷款类型	利率	坏账比率
个人	0.140	0.10
汽车	0.130	0.07
住房	0.120	0.03
农场	0.125	0.05
商业	0.100	0.02

坏账不可收回且不产生利息收入.

为了与其他金融机构竞争, 要求银行把至少 40% 的资金分配给农场和商业贷款. 为扶持当地的住房产业, 住房贷款至少要等于个人、汽车和住房贷款总额的 50%. 银行还有一项明确的政策, 不允许坏账的总比例超过全部贷款的 4%.

数学模型 问题是寻求确定每一种贷款的数额 (单位: 100 万美元), 因此给出下面变量的定义:

$$x_1 = \text{个人贷款}$$
$$x_2 = \text{汽车贷款}$$
$$x_3 = \text{住房贷款}$$
$$x_4 = \text{农场贷款}$$
$$x_5 = \text{商业贷款}$$

Thriftem 银行的目标是使得净收益 (即利息收入与坏账损失之差) 达到最大. 只有良性的贷款才会产生利息收入. 由于 10% 个人贷款属于收不回的坏账, 因此, 银行只能对其中 90% 的贷款收取利息, 即收取原贷款额 x_1 中 $0.9x_1$ 的 14% 的利息. 对其余 4 种贷款, 同理可得

$$\begin{aligned}\text{总利息} &= 0.14(0.9x_1) + 0.13(0.93x_2) + 0.12(0.97x_3) + 0.125(0.95x_4) + 0.1(0.98x_5) \\ &= 0.126x_1 + 0.120\,9x_2 + 0.116\,4x_3 + 0.118\,75x_4 + 0.098x_5\end{aligned}$$

我们还有

$$\text{坏账} = 0.1x_1 + 0.07x_2 + 0.03x_3 + 0.05x_4 + 0.02x_5$$

因此, 目标函数表达为

$$\begin{aligned}\max \quad z &= \text{总利息} - \text{坏账} \\ &= (0.126x_1 + 0.120\,9x_2 + 0.116\,4x_3 + 0.118\,75x_4 + 0.098x_5) \\ &\quad -(0.1x_1 + 0.07x_2 + 0.03x_3 + 0.05x_4 + 0.02x_5) \\ &= 0.026x_1 + 0.050\,9x_2 + 0.086\,4x_3 + 0.068\,75x_4 + 0.078x_5\end{aligned}$$

这个问题有 5 个约束:

(1) *资金总额应该不超过 1 200 万美元:*

$$x_1 + x_2 + x_3 + x_4 + x_5 \leqslant 12$$

(2) *农场和商业贷款至少等于总贷款的 40%:*

$$x_4 + x_5 \geqslant 0.4(x_1 + x_2 + x_3 + x_4 + x_5)$$

或

$$0.4x_1 + 0.4x_2 + 0.4x_3 - 0.6x_4 - 0.6x_5 \leqslant 0$$

(3) *住房贷款应该至少等于个人、汽车和房屋贷款的 50%:*

$$x_3 \geqslant 0.5(x_1 + x_2 + x_3)$$

或

$$0.5x_1 + 0.5x_2 - 0.5x_3 \leqslant 0$$

(4) *坏账比例不应超过总贷款的 4%:*

$$0.1x_1 + 0.07x_2 + 0.03x_3 + 0.05x_4 + 0.02x_5 \leqslant 0.04(x_1 + x_2 + x_3 + x_4 + x_5)$$

或

$$0.06x_1 + 0.03x_2 - 0.01x_3 + 0.01x_4 - 0.02x_5 \leqslant 0$$

(5) *非负性约束:*

$$x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, x_4 \geqslant 0, x_5 \geqslant 0$$

在前面的公式中, 一个不十分明显的假定是, 所有的贷款都是差不多同时发放的. 这一假设允许我们不考虑分配给不同贷款资金的时间价值差.

解 最优解为 $z = 0.996\,48, x_1 = 0, x_2 = 0, x_3 = 7.2, x_4 = 0, x_5 = 4.8$

评注

(1) 你可能觉得奇怪, 为什么不将第 2 个约束的右端项定义为 0.4×12 来替代 $0.4(x_1 + x_2 + x_3 + x_4 + x_5)$. 毕竟, 它似乎是合理的, 因为银行希望的总贷款额为 1 200 万美元. 答案是第 2 种用法并没有“剥夺”模型的这种可能性. 如果最优解需要全部 1 200 万美元的贷款额, 给出的约束将允许这样. 但此处不使用 0.4×12 有两个重要的原因: ① 如果在模型中其他约束限制了全部贷款额在 1 200 万美元以

下 (例如, 银行可以为不同的贷款设置上限), 那么选择 0.4×12 则可能导致不可行或不正确的解; ② 如果打算改变可用资金上限 (如从 1 200 万美元变到 1 300 万美元) 而尝试最优解的效果, 可能你会忘记将 0.4×12 改为 0.4×13, 在这种情况下, 得到的结果将是不正确的. 对于第 4 个约束的右端项也出于类似的道理.

(2) 最优解要求分配所有 1 200 万美元中的 720 万美元用于住房贷款, 480 万美元用于商业贷款. 没有发放其他类型的贷款. 投资收益的计算如下:

$$\text{收益率} = \frac{z}{12} = \frac{0.996\ 48}{12} = 0.080\ 34$$

这表明, 组合投资的年收益率是 8.034%, 它低于最佳净利率 (住房贷款为 0.086 4), 你会感到惊讶, 为什么最优解达不到这个值. 答案是银行规定了农场和商业贷款至少占总贷款的 40%(第 2 个约束), 这迫使解以较低的净利率 0.078 分配 480 万美元给商业贷款, 因此降低了总利率, 为 $\frac{0.086\ 4 \times 7.2 + 0.078 \times 4.8}{12} = 0.080\ 34$. 事实上, 如果去掉第 2 个约束, 最优解将以较高的利率 8.46% 把全部资金分配给住房贷款.

习题 2.3C

1. Fox 建筑公司正在考虑未来的 4 年内 6 项可能的建筑工程. 工程的期望 (现值) 收益和现金支出额已知如下. Fox 公司可以部分或全部承担任何一项. 工程承担的部分将按比例分配收益和现金支出额. (单位: 1 000 美元)

工程	现金支出额				收益
	第 1 年	第 2 年	第 3 年	第 4 年	
1	10.5	14.4	2.2	2.4	32.40
2	8.3	12.6	9.5	3.1	35.80
3	10.2	14.2	5.6	4.2	17.75
4	7.2	10.5	7.5	5.0	14.80
5	12.3	10.1	8.3	6.3	18.20
6	9.2	7.8	6.9	5.1	12.35
可用资金	60.0	70.0	35.0	20.0	

(a) 将该问题表述成一个线性规划, 确定最优的工程混合, 使得总收益最大. 忽略资金的时间价值.

(b) 假定如果承担工程 2 的一部分, 则至少要承担工程 6 相等的部分. 修改模型的公式, 并求新的最优解.

(c) 在原模型中, 假定到年底剩下资金可用于下一年. 求新的最优解, 并确定每一年从前一年"借多少钱". 为了简化起见, 忽略资金的时间价值.

(d) 假定在原始模型中, 任何年度可用资金可以通过借款的方式超出, 借款来自公司内其他金融项目. 忽略资金的时间价值, 重建线性规划模型, 并求最优解. 新的结果需要在任何一年借款吗? 如果需要, 借来资金的收益率是多少?

***2.** 投资者 Doe 有 10 000 美元, 打算在 4 个项目上投资. 下表给出了 4 项投资的现金流.

项目	某年开始的现金流 (1 000 美元)				
	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年
1	−1.00	0.50	0.30	1.80	1.20
2	−1.00	0.60	0.20	1.50	1.30
3	0.00	−1.00	0.80	1.90	0.80
4	−1.00	0.40	0.60	1.80	0.95

表中的数据解释如下: 对于项目 1, 在第 1 年初投资的 1 美元, 将在第 2 年初产生 0.50 美元收益, 第 3 年初为 0.30 美元, 第 4 年初为 1.80 美元, 第 5 年初为 1.20 美元. 其余各项有相同的解释. 有一项为 0.00, 表明该年没有发生交易. Doe 还有另外一项投资选择, 银行存款的年利率为 6.5%. 所有积累的资金在年底可用于下一年的再投资. 将上述问题建立成一个线性规划, 确定对于各投资机会的最优资金分配.

3. HiRise 建筑公司可以投标两个一年期的项目. 下表提供了两个项目每季度的现金流 (100 万美元).

项目	在下列日期的现金流				
	1/1/08	4/1/08	7/1/08	10/1/08	12/31/08
I	−1.0	−3.1	−1.5	1.8	5.0
II	−3.0	−2.5	1.5	1.8	2.8

HiRise 在每季度的开始有 100 万美元的现金, 并且最多可以借 100 万美元, 其票面年利率为 10%. 任何借款必须在季度末返还. 剩余的现金可以获得季度的利息, 其票面年利率为 8%. 一个季度末的净积累可作为下季度的投资.

(a) 假定允许 HiRise 部分或全部参与两个项目. 求参与水平, 使得 12/31/2008 的净现金积累量达到最大.

(b) 是否有可能在任何一季度既有借款且同时又有剩余资金? 试解释之.

4. 预料到孩子上学需要一笔很大的费用, 一对夫妇开始了一项年度投资计划, 其时间是从他们孩子的第 8 个生日一直持续到孩子的第 18 个生日. 这对夫妇估计, 在每年年初, 他们能够投入下列数量的资金 (美元):

年	1	2	3	4	5	6	7	8	9	10
数量	2 000	2 000	2 500	2 500	3 000	3 500	3 500	4 000	4 000	5 000

为避免不愉快的意外, 他们打算安全地投资在下列选项: 保险储蓄, 其年增长率为 7.5%; 6 年期公债, 其年增长率为 7.9%, 并且当前的市场价格等于面值的 98%; 9 年期市政公债, 其年增长率为 8.5%, 而且当前的市场价格是面值的 1.02 倍. 这对夫妇应该如何投资呢?

***5.** 一位商业主管在两项计划中有投资选择权: 计划 A 保证每一美元的投资在 1 年后能赚得 0.70 美元, 而计划 B 保证每一美元的投资在 2 年后能赚得 2 美元. 对于计划 A, 可以按年制订投资规划, 而对于计划 B, 只允许以 2 年为周期制订投资规划. 主管应如何投资 100 000 美元, 使得在 3 年末的收入达到最大?

6. 一赌徒在玩一种需要在 4 项选择中划分赌金的游戏. 游戏有 3 个结果. 下表给出了针对游戏的不同选项, 每美元相应的获得与损失的钱数.

结果	投注下列选项每美元的收益			
	1	2	3	4
1	−3	4	−7	15
2	5	−3	9	4
3	3	−9	10	−8

该赌徒共有 500 美元, 可以只玩一次. 游戏的确定结果是事先不知道的. 因为这种不确定性, 赌徒的策略是由 3 个结果产生的最小收益达到最大. 该赌徒在 4 种选择中应该如何分配这 500 美元呢? (提示: 赌徒的净收益可能是正的、零或负的).

7. (Lewis, 1996) 一户家庭, 每月会收到一些账单, 这些账单有按月支付的 (例如, 水电费和家庭按揭贷款), 有按季度支付的 (如, 估计的纳税), 有按半年支付的 (例如, 保险), 或按年支付的 (例如, 订报费和会费). 下表提供了下一年每月的账单.

月	1	2	3	4	5	6	7	8	9	10	11	12	总额
$	800	1 200	400	700	600	900	1 500	1 000	900	1 100	1 300	1 600	12 000

为支付这些费用, 家庭每月留出 1 000 美元, 它是总额除以 12 的平均值. 如果将钱放在定期储蓄账户中, 倘若存款时间超过 1 个月以上, 可赚得 4% 的年利率. 银行也同时提供 3 个月和 6 个月存款方式, 其年利率分别是 5.5% 和 7%. 设计一个 12 个月的投资时间表, 使该家庭全年的总收益最大. 说明达到可行解必需的任何假定或要求.

2.3.4 生产计划和库存控制

大量地应用于生产和库存控制的线性规划, 应用范围从简单的加工能力配置以满足需求, 到利用库存来"平抑"由于计划期内需求的突然变化而带来的影响, 以及采取雇工和解雇策略来回应劳动力需求的变化等比较复杂的情况. 本节介绍 3 个例子. 第 1 个例子介绍在单一周期内使用公共生产设备以满足需求的产品安排. 第 2 个例子介绍在多周期生产系统中如何使用库存来满足未来的需求. 第 3 个例子是关于组合运用库存、雇工与解雇手段, 以"平滑"多周期计划期内上下浮动的需求生产安排.

例 2.3-4 (单周期生产模型)

为了满足冬季市场需求, 某服装公司正在加工皮制外衣、鹅绒外套、保暖裤和手套. 所有产品由 4 个不同的车间生产: 剪裁、保暖处理、缝纫和包装. 服装公司已收到其他公司的产品订单. 合同规定对于未按时交货的订单产品将予以惩罚. 下表提供了本问题的相关数据.

车间	每件产品的生产时间 (小时)				生产能力 (小时)
	皮制外衣	鹅绒外套	保暖裤	手套	
剪裁	0.30	0.30	0.25	0.15	1 000
保暖	0.25	0.35	0.30	0.10	1 000
缝纫	0.45	0.50	0.40	0.22	1 000
包装	0.15	0.15	0.1	0.05	1 000
需求	800	750	600	500	
单位利润	$30	$40	$20	$10	
单位惩罚	$15	$20	$10	$8	

试为公司设计最优的生产计划.

数学模型 变量的定义很简单. 令

$$x_1 = \text{皮制外衣的数量}$$

$$x_2 = \text{鹅绒外衣的数量}$$

$$x_3 = \text{保暖裤的数量}$$

$$x_4 = \text{手套的数量}$$

当需求不满足时, 公司会被处罚. 这意味着, 问题的目标是极大化净收入, 定义为

$$\text{净收入} = \text{总利润} - \text{总惩罚}$$

总利润很容易地表示为 $30x_1 + 40x_2 + 20x_3 + 10x_4$. 总惩罚是短缺量的函数 (= 需求量 − 每种产品的供应量). 这些短缺量可以由下列需求上限来确定:

$$x_1 \leqslant 800,\ x_2 \leqslant 750,\ x_3 \leqslant 600,\ x_4 \leqslant 500$$

如果需求约束满足严格的不等式, 则没有满足相应的需求. 例如, 如果生产了 650 件皮夹克, 则 $x_1 = 650$, 这样就得到了皮夹克的短缺量 $800 - 650 = 150$. 可以定义新的非负变量用代数形式表示任何产品的短缺量, 即

$$s_j = \text{产品 } j \text{ 的短缺数量},\quad j = 1, 2, 3, 4$$

在这种情况下, 需求约束可以被写成

$$x_1 + s_1 = 800,\ x_2 + s_2 = 750,\ x_3 + s_3 = 600,\ x_4 + s_4 = 500$$

$$x_j \geqslant 0,\quad s_j \geqslant 0,\quad j = 1, 2, 3, 4$$

现在用 $15s_1 + 20s_2 + 10s_3 + 8s_4$ 来计算出短缺惩罚的费用. 因此, 目标函数可以写成

$$\max\ z = 30x_1 + 40x_2 + 20x_3 + 10x_4 - (15s_1 + 20s_2 + 10s_3 + 8s_4)$$

为得到完整的模型, 还有些约束是表示生产能力的限制, 即

$$0.30x_1 + 0.30x_2 + 0.25x_3 + 0.15x_4 \leqslant 1\,000\ \text{(剪裁)}$$

$$0.25x_1 + 0.35x_2 + 0.30x_3 + 0.10x_4 \leqslant 1\,000\ \text{(保暖)}$$

$$0.45x_1 + 0.50x_2 + 0.40x_3 + 0.22x_4 \leqslant 1\,000\ \text{(缝纫)}$$

$$0.15x_1 + 0.15x_2 + 0.10x_3 + 0.05x_4 \leqslant 1\,000\ \text{(包装)}$$

因此完整的模型为

$$\max\ z = 30x_1 + 40x_2 + 20x_3 + 10x_4 - (15s_1 + 20s_2 + 10s_3 + 8s_4)$$

$$\text{s.t.}\ \ 0.30x_1 + 0.30x_2 + 0.25x_3 + 0.15x_4 \leqslant 1\,000$$

$$0.25x_1 + 0.35x_2 + 0.30x_3 + 0.10x_4 \leqslant 1\,000$$

$$0.45x_1 + 0.50x_2 + 0.40x_3 + 0.22x_4 \leqslant 1\ 000$$
$$0.15x_1 + 0.15x_2 + 0.10x_3 + 0.05x_4 \leqslant 1\ 000$$
$$x_1 + s_1 = 800,\ x_2 + s_2 = 750,\ x_3 + s_3 = 600,\ x_4 + s_4 = 500$$
$$x_j \geqslant 0,\quad s_j \geqslant 0,\quad j = 1, 2, 3, 4$$

解 最优解是 $z = 64\ 625$ 美元, $x_1 = 800$, $x_2 = 750$, $x_3 = 387.5$, $x_4 = 500$, $s_1 = s_2 = s_4 = 0$, $s_3 = 212.5$. 结果满足两种外衣和手套的所有需求. 保暖裤短缺 213 条 (由 212.5 取整得到), 惩罚费用是 $213 \times 10 = 2\ 130$ 美元.

例 2.3-5 (多周期生产–库存模型)

Acme 制造公司已签订了未来的 6 个月提供房屋窗户的合同. 每月的需求量分别是 100, 250, 190, 140, 220, 110 扇. 每扇窗户的生产成本与劳动力、原材料和水电费用有关, 每个月都不同. Acme 公司估计在未来的 6 个月, 每扇窗户的生产成本分别是 50 美元、45 美元、55 美元、48 美元、52 美元和 50 美元. 为了利用生产成本变动的有利条件, Acme 公司可以选择生产多于某个月的需求, 而保存剩余的部分为后面月份交货. 然而, 这将导致每月每扇窗户有 8 美元的存储成本. 建立一个线性规划, 确定最优的产品生产安排.

数学模型 本问题的变量包括月生产量和月底的库存量. 对于 $i = 1, 2, \cdots, 6$, 令

$$x_i = \text{第 } i \text{ 个月的生产数量}$$
$$I_i = \text{第 } i \text{ 个月月底的库存数}$$

这些变量与未来 6 个月范围内月需求之间的关系如图 2.5 所示. 系统以零值开始, 这意味着 $I_0 = 0$.

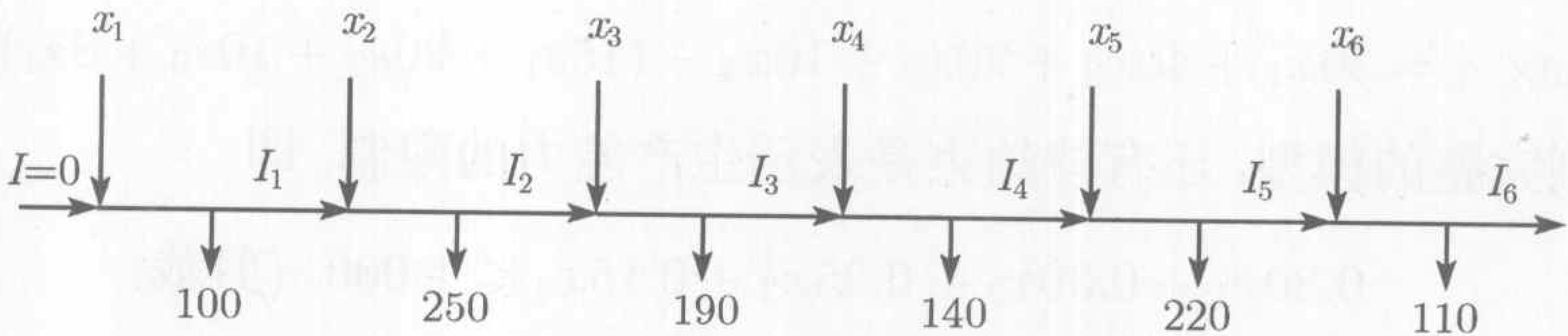


图 2.5 生产–库存系统的示意图

目标函数是求生产成本与月末库存成本之和的最小值. 这里我们有

$$\text{总生产成本} = 50x_1 + 45x_2 + 55x_3 + 48x_4 + 52x_5 + 50x_6$$
$$\text{总库存成本} = 8(I_1 + I_2 + I_3 + I_4 + I_5 + I_6)$$

因此目标函数是

$$\min\ z = 50x_1 + 45x_2 + 55x_3 + 48x_4 + 52x_5 + 50x_6$$

$$+8(I_1+I_2+I_3+I_4+I_5+I_6)$$

问题的约束可以由图 2.5 直接得到. 对于每个周期, 我们有下列平衡方程:

月初库存 + 生产量 − 月末库存 = 需求

按月写成数学表达式如下:

$$\begin{aligned}
I_0+x_1-I_1 &= 100 \ (\text{第 1 个月})\\
I_1+x_2-I_2 &= 250 \ (\text{第 2 个月})\\
I_2+x_3-I_3 &= 190 \ (\text{第 3 个月})\\
I_3+x_4-I_4 &= 140 \ (\text{第 4 个月})\\
I_4+x_5-I_5 &= 220 \ (\text{第 5 个月})\\
I_5+x_6-I_6 &= 110 \ (\text{第 6 个月})\\
x_i,\ \ I_i \geqslant 0, &\quad i=1,2,\cdots,6\\
I_0 &= 0
\end{aligned}$$

对于本问题, $I_0=0$, 因为问题开始时没有初始库存. 还有, 在任何最优解中, 最终库存 I_6 也将是 0, 因为交货完毕还有存货是不符合逻辑的, 它只能导致额外的库存成本却没有任何实际意义.

现在给出完整的数学模型如下:

$$\begin{aligned}
\min \quad & z=50x_1+45x_2+55x_3+48x_4+52x_5+50x_6\\
& \qquad +8(I_1+I_2+I_3+I_4+I_5+I_6)\\
\text{s.t.} \quad & x_1-I_1=100 \ (\text{第 1 个月})\\
& I_1+x_2-I_2=250 \ (\text{第 2 个月})\\
& I_2+x_3-I_3=190 \ (\text{第 3 个月})\\
& I_3+x_4-I_4=140 \ (\text{第 4 个月})\\
& I_4+x_5-I_5=220 \ (\text{第 5 个月})\\
& I_5+x_6-I_6=110 \ (\text{第 6 个月})\\
& x_i,\ \ I_i \geqslant 0, \ \ i=1,2,\cdots,6
\end{aligned}$$

解 图 2.6 概括了问题的最优解. 它表明, 每月的需求由每月的生产直接满足, 除了第 2 个月生产 440 扇, 用来满足第 2 个月和第 3 个月的需求. 相应的总费用是 49 980 美元.

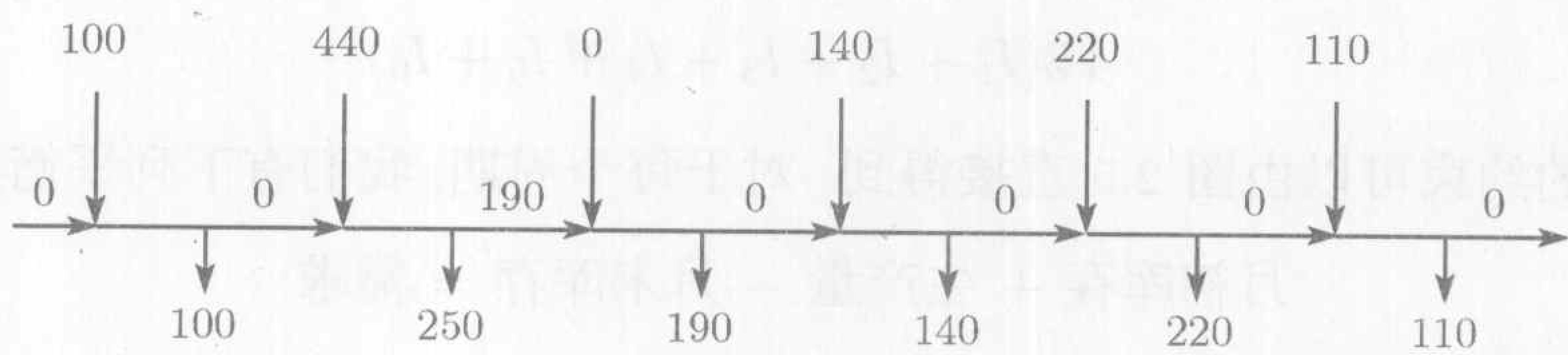


图 2.6　生产与库存的最优解

例 2.3-6 (多周期生产平滑模型)

一家公司将在未来的 4 个月 (3 月、4 月、5 月和 6 月) 生产某种产品. 每月的需求量分别为 520 件、720 件、520 件和 620 件. 公司有 10 位正式工, 但是, 如果需要的话, 可通过雇用和解雇临时工来适应上下变动的生产需求. 在任何一个月, 雇用和解雇的额外成本分别是每名工人 200 美元和 400 美元. 一名正式工人每月能生产 12 件, 而一名临时工由于缺乏相应的经验, 每月只能生产 10 件. 在任何一个月, 公司的生产可以多于需求, 并将过剩的产品转到后面的某个月份, 每件每月的库存成本为 50 美元. 试为公司这 4 个月的计划设计一种最优的雇用/解雇策略.

数学模型　这个模型乍看起来与例 2.3-5 类似, 每月有它的生产、需求和月末库存. 但有两处例外: (1) 需要考虑正式工与临时工之间的关系; (2) 需要考虑每月的雇用和解雇成本.

因为 10 名正式工不能够被解雇, 因此, 可以分别从每个月的需求减去他们的产量来表示他们的影响. 即便有其余的需求, 也是通过雇用和解雇临时工来满足. 从模型的观点来看, 每月的净余需求是

$$\begin{aligned}
&3\text{ 月份的需求} = 520 - 12 \times 10 = 400\text{ 件}\\
&4\text{ 月份的需求} = 720 - 12 \times 10 = 600\text{ 件}\\
&5\text{ 月份的需求} = 520 - 12 \times 10 = 400\text{ 件}\\
&6\text{ 月份的需求} = 620 - 12 \times 10 = 500\text{ 件}
\end{aligned}$$

对于 $i = 1, 2, 3, 4$, 模型的变量可以定义为

$$\begin{aligned}
x_i &= \text{在任何雇用或解雇后, 第 } i \text{ 个月初临时工的净人数}\\
S_i &= \text{第 } i \text{ 个月初, 雇用或解雇临时工的数量}\\
I_i &= \text{第 } i \text{ 个月月末库存产品件数}
\end{aligned}$$

由定义, 变量 x_i 和 I_i 必须假定是非负的. 但是, 变量 S_i 可以是正的 (当雇用新的临时工时); 可以是负的 (当解雇临时工时); 可以是零 (如果没有雇用和解雇发生). 其结果是, 该变量必须是*无符号限制*. 这是本章使用无符号限制变量的第一个例子. 正像我们不久将看到的那样, 需要特殊的替代才能允许实现模型的雇用与解雇.

目标是极小化雇用与解雇成本, 以及从本月到下月的库存成本三者之和. 库存成本的处理类似于例 2.3-5 给出的情况, 即

$$\text{库存成本} = 50\,(I_1 + I_2 + I_3)$$

(注意, 在最优解中, $I_4 = 0$.) 雇用和解雇成本有些复杂. 我们知道在任何最优解中, 至少 40 名临时工 ($= \frac{400}{10}$) 必须在 3 月初被雇用, 以适应该月的需求. 然而, 并不将这种情况作为特殊情况处理, 我们将留给最优化过程来自动处理. 因此, 已知雇用和解雇临时工的成本分别为 200 美元和 400 美元, 我们有

$$\begin{aligned}\text{雇用和解雇的成本} \;=\; & 200 \times \text{在 3 月、4 月、5 月和 6 月初雇用临时工的数量}\\ & +400 \times \text{在 3 月、4 月、5 月和 6 月初解雇临时工的数量}\end{aligned}$$

用数学表达式翻译这个方程, 首先需要建立约束条件.

模型的约束涉及库存、雇用和解雇. 首先建立库存约束. 定义 x_i 为第 i 个月可用的临时工的数量, 并且已知临时工每个月的生产能力是 10 件, 所以在第 i 个月的生产量是 $10x_i$. 因此, 库存约束为

$$\begin{aligned}
10x_1 &= 400 + I_1 & &(3\text{ 月})\\
I_1 + 10x_2 &= 600 + I_2 & &(4\text{ 月})\\
I_2 + 10x_3 &= 400 + I_3 & &(5\text{ 月})\\
I_3 + 10x_4 &= 500 & &(6\text{ 月})\\
x_1, x_2, x_3, x_4 &\geqslant 0, & &I_1, I_2, I_3 \geqslant 0
\end{aligned}$$

接下来, 建立关于雇用和解雇的约束. 首先注意到, 临时工在 3 月初有 x_1 名工人. 然后在 4 月初, 劳动力人数由 x_1 调整 (增加或减少)S_2 得到 x_2. 对于 x_3 和 x_4 用类似的处理方法, 得到下列方程:

$$\begin{aligned}
x_1 &= \quad\;\; S_1\\
x_2 &= x_1 + S_2\\
x_3 &= x_2 + S_3\\
x_4 &= x_3 + S_4\\
& S_1, S_2, S_3, S_4\text{无符号限制}\\
& x_1, x_2, x_3, x_4 \geqslant 0
\end{aligned}$$

变量 S_1, S_2, S_3, S_4 当它们是严格正时表示雇用, 当它们是严格负时表示解雇. 然而, 这种“定性的”信息并不能用在数学表达式中. 而需要作下列替代:

$$S_i = S_i^- - S_i^+, \quad \text{其中} \quad S_i^-, S_i^+ \geqslant 0$$

无限制变量 S_i 现在是两个非负变量 S_i^- 和 S_i^+ 的差. 可以认为 S_i^- 是雇用临时工的数量, S_i^+ 是解雇临时工的数量. 例如, 如果 $S_i^- = 5$ 和 $S_i^+ = 0$, 则 $S_i = 5 - 0 = +5$,

它表示雇用. 如果 $S_i^- = 0$, $S_i^+ = 7$, 则 $S_i = 0 - 7 = -7$, 它表示解雇. 在第 1 种情况下, 相应的雇用成本是 $200S_i^- = 200 \times 5 = 1\,000$ 美元, 在第 2 种情况下, 相应的解雇成本是 $400S_i^+ = 400 \times 7 = 2\,800$ 美元. 这一思想是构造目标函数的基础.

首先需要说明重要的一点是: 如果 S_i^- 和 S_i^+ 同时为正, 那么这表示什么? 其答案是, 这是不可能发生的, 因为它蕴涵着其结果要求一个月内既要雇用工人, 同时又要解雇工人. 有趣的是, 线性规划理论 (见第 13 章) 告诉我们 S_i^- 和 S_i^+ 不可能同时为正, 这个结果与我们的直觉是一致的.

现在可以写出雇用和解雇成本如下:

$$\text{雇用成本} = 200\left(S_1^- + S_2^- + S_3^- + S_4^-\right)$$
$$\text{解雇成本} = 400\left(S_1^+ + S_2^+ + S_3^+ + S_4^+\right)$$

完整的模型是

$$\begin{aligned}
\min\ z = {} & 50\left(I_1 + I_2 + I_3 + I_4\right) + 200\left(S_1^- + S_2^- + S_3^- + S_4^-\right) \\
& + 400\left(S_1^+ + S_2^+ + S_3^+ + S_4^+\right) \\
\text{s.t.}\quad & 10x_1 = 400 + I_1 \\
& I_1 + 10x_2 = 600 + I_2 \\
& I_2 + 10x_3 = 400 + I_3 \\
& I_3 + 10x_4 = 500 \\
& x_1 = S_1^- - S_1^+ \\
& x_2 = x_1 + S_2^- - S_2^+ \\
& x_3 = x_2 + S_3^- - S_3^+ \\
& x_4 = x_3 + S_4^- - S_4^+ \\
& S_1^-, S_1^+, S_2^-, S_2^+, S_3^-, S_3^+, S_4^-, S_4^+ \geqslant 0 \\
& x_1, x_2, x_3, x_4 \geqslant 0 \\
& I_1, I_2, I_3 \geqslant 0
\end{aligned}$$

解 最优解是 $z = 19\,500$ 美元, $x_1 = 50$, $x_2 = 50$, $x_3 = 45$, $x_4 = 45$, $S_1^- = 50$, $S_3^+ = 5$, $I_1 = 100$, $I_3 = 50$. 其余的所有变量均为 0. 结果表明, 在 3 月份雇用临时工 50 名 ($S_1^- = 50$), 并且保持稳定的劳动力到 5 月份, 在 5 月份解雇临时工 5 名 ($S_3^+ = 5$). 其他的时间不再雇用和解雇工直到 6 月底, 那时, 解雇所有的临时工. 这个结果表明 4 月份有上月的 100 件库存, 在 6 月份有上个月的 50 件库存.

习题 2.3D

1. Toolco 公司与 AutoMate 公司签订了为其自助廉价商店供应扳手和凿子的合同. AutoMate 每周的需求至少是 1 500 个扳手和 1 200 个凿子. 按照现行的单班生产能力, Toolco 不能完成所有的订单数, 必须加班生产, 也可以转包给其他工具厂完成合同. 这样会增加每件产品的生产成本, 如下表所示. 市场需求规定凿子与扳手的比例至少是 2:1.

工具	生产类型	每周生产的范围 (数量)	单位费用 (美元)
扳手	正常	0~550	2.00
	加班	551~800	2.80
	转包	801~∞	3.00
凿子	正常	0~620	2.10
	加班	621~900	3.20
	转包	901~∞	4.20

(a) 将本问题写成一个线性规划, 确定每种工具最优的生产进度.

(b) 陈述如下与模型合理性相联系的事实: 生产成本函数包含递增单位成本.

2. 在 3 台机器上相继加工 4 种产品. 下表给出了本问题相关的数据.

机器	每小时费用 (美元)	每件产品的制造时间 (小时)				能力 (小时)
		产品 1	产品 2	产品 3	产品 4	
1	10	2	3	4	2	500
2	5	3	2	1	2	380
3	4	7	3	2	1	450
单位售价 (美元)		75	70	55	45	

建立本问题的线性规划模型, 并求最优解.

***3.** 某个制造商使用原料 A 和 B 生产某种产品的 3 种型号 I、Ⅱ 和 Ⅲ. 下表给出了问题的数据.

原料	每件产品对原料的需求			可用量
	I	Ⅱ	Ⅲ	
A	2	3	5	4 000
B	4	2	7	6 000
最小需求量	200	200	150	
单位利润 (美元)	30	20	50	

每件型号 I 产品的劳动时间是型号 Ⅱ 的 2 倍, 是型号 Ⅲ 的 3 倍. 该厂的全部劳动力能够生产相当于 1 500 件型号 I 的产品. 市场对于 3 种不同型号产品需求的特定比例是 3:2:5. 将问题建立成一个线性规划, 并求最优解.

4. All-Flavors 冰淇淋店在整个夏季的 3 个月中 (6 月、7 月和 8 月) 对于冰淇淋的需求估计分别是 500, 600 和 400 箱. 有两个冰淇淋批发商 1 和 2 向 All-Flavors 供货. 虽然这两个供应商的冰淇淋的风味不同, 但可以互相交换. 任何一个供应商能够提供冰淇淋的最大箱数是每月 400 箱. 还有, 这两个供应商的供货价格是按照下面的时间表逐月变化的.

	一个月中每箱冰淇淋的价格 (美元)		
	6 月	7 月	8 月
供应商 1	100	110	120
供应商 2	115	108	125

为了利用价格波动, All-Flavors 购买的冰淇淋可以多于某个月的需求, 并且储存剩余的部分以满足以后各月的需求. 冷藏一箱冰淇淋的成本是每月 5 美元. 实际上可以假定, 冷藏成本是当月持有冰淇淋平均箱数的函数. 建立一个从这两个供应商购买冰淇淋的最优采购计划.

5. 对某种产品在未来 4 个季度的需求量分别是 300, 400, 450, 250 件. 每件的价格在第 1 季度以 20 美元开始, 其随后的每个季度增加 2 美元. 供应商在任一季度最多可以提供产品 400 件. 尽管我们可以利用前面季度的低价优势, 但它会导致每季度每件 3.5 美元的储存成本. 还有, 从一个季度储存到下一季度的最大件数不能超过 100. 试为该产品建立一个最优采购计划以满足需求.

6. 某公司已签订了在未来的 6, 7, 8 月份生产 A, B 两种产品的合同. 总的生产能力 (用小时表示) 每月不同. 下表给出了本问题的基本数据:

	6 月	7 月	8 月
对 A 的需求 (件)	500	5 000	750
对 B 的需求 (件)	1 000	1 200	1 200
生产能力 (小时)	3 000	3 500	3 000

对于产品 A 和 B, 每小时的生产率分别是 1.25 件和 1 件. 所有的需求必须被满足. 然而, 后面月份的需求可以从前面的月份的生产来填补. 对于任何从本月转到下一月的产品, 每件产品 A 和产品 B 每月分别有 0.90 美元和 0.75 美元的储存成本. A, B 两种产品的单位生产成本分别是 30 美元和 28 美元. 试为两种产品确定最优的生产计划安排.

***7.** 某产品的制造过程由前后两道工序 I 和工序 Ⅱ 组成. 下表提供了在未来的 6, 7, 8 月份的相关数据:

	6 月	7 月	8 月
成品的需求 (件)	500	450	600
工序 I 的能力 (小时)	800	700	550
工序 Ⅱ 的能力 (小时)	1 000	850	700

生产一件产品在工序 I 上花 0.6 小时, 在工序 Ⅱ 另外花 0.8 小时. 在任何一个月过剩的产品, 无论是半成品 (工序 I) 还是成品 (工序 Ⅱ), 都允许在后面的月份中使用. 相应的储存成本是每件每月 0.20 美元和 0.40 美元. 生产成本随工序和月份变化. 对于工序 I, 单位生产成本在 6, 7, 8 月份分别为 10 美元、12 美元和 11 美元. 对于工序 Ⅱ, 相应的单位生产费用分别为 15 美元、18 美元和 16 美元. 确定这两道工序在未来的 3 个月内最优的生产进度安排.

8. 在两台机器上相继加工两种产品. 每台机器的工作时间是每天 8 小时, 必要时, 可以通过加班的方式增加 4 小时, 其附加成本是每小时 100 美元. 下表给出了两台机器的生产率和两种产品的单位价格. 确定最优的生产计划安排, 并根据需要给出所用的加班时间.

	生产率 (件/小时)	
	产品 1	产品 2
机器 1	5	5
机器 2	8	4
单位价格 (美元)	110	118

2.3.5 混合与精炼

一些线性规划应用涉及混合不同的输入原料来生产满足特定规格的产品, 以使得成本最低或利润最大. 输入原料可能是矿石、金属废料、化学药品或原油, 而输出产品可能是各种不同等级的金属制锭、油漆或汽油. 本节介绍一种 (简单的) 石油精炼模型, 其过程是从蒸馏原油开始, 以生产中间的汽油存料, 然后再混合这些存料生产出最终的汽油. 最终产品必须满足特定的质量规格 (如辛烷等级). 此外, 蒸馏能力和需求限制可以直接影响不同等级汽油的生产水平. 模型的一个目标是确定最终产品的最优混合, 使得某种利润函数达到极大值. 在某些情形下, 目标也可以是求成本函数的最小值.

例 2.3-7 (原油精炼和汽油混合)

Shale 石油公司位于 Aruba 岛上, 它具有日产 1 500 000 桶原油的能力. 精炼厂的最终产品是含有不同辛烷值 (ON) 的 3 种类型无铅汽油: 普通汽油 ON= 87, 优质汽油 ON= 89, 超优汽油 ON= 92. 精炼过程包含 3 个阶段: (1) 使用蒸馏塔来生产存料 (ON= 82), 其生产率为每桶原油生产 0.2 桶存料; (2) 使用裂化装置把蒸馏塔得到的部分存料生产为汽油存料 (ON= 98), 其生产率为每桶存料生产 0.5 桶汽油存料; (3) 使用混合装置来混合裂化装置产的汽油存料与蒸馏塔产的存料. 公司估计, 3 种汽油的每桶净利润分别为 6.70 美元、7.20 美元和 8.10 美元. 裂化装置的输入能力是每天 200 000 桶存料. 对于普通汽油、优质汽油和超优汽油的需求上限分别为每天 50 000 桶、30 000 桶和 40 000 桶. 为精炼厂建立一个模型以确定最优生产进度安排.

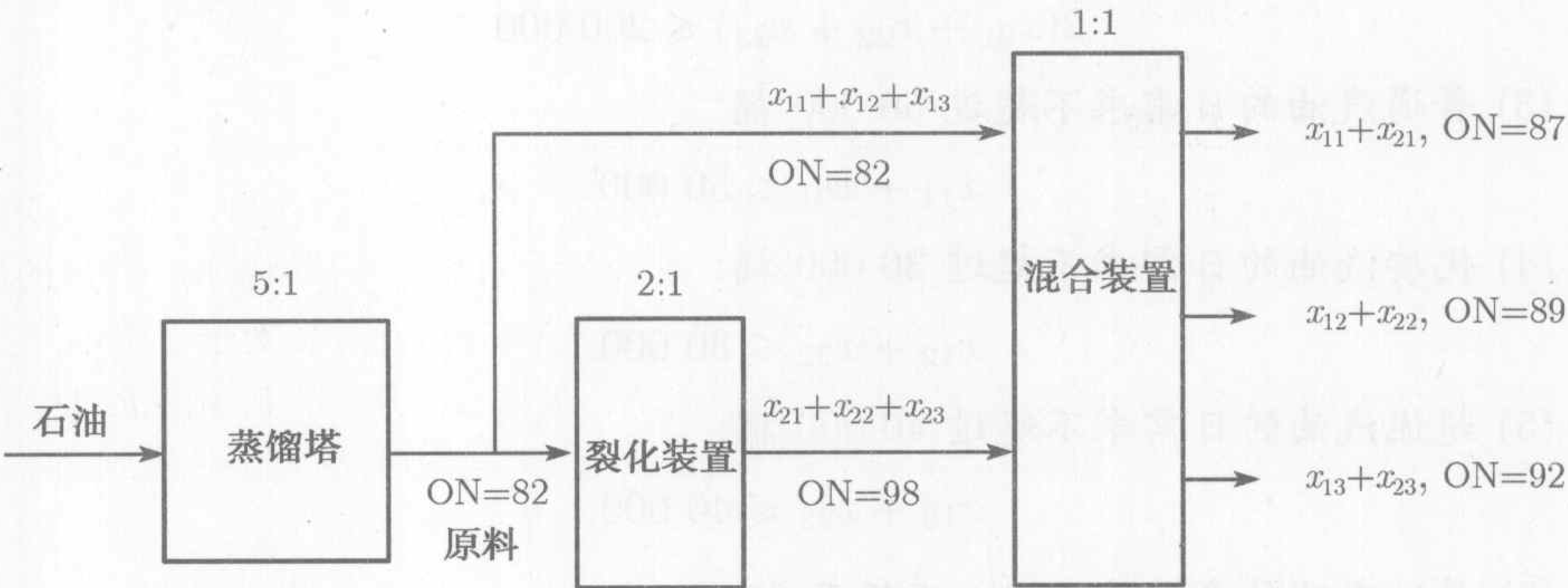


图 2.7 精炼厂问题的产品流

数学模型 图 2.7 概述了模型的要素. 按照进入混合装置的 2 条输入流 (存料与裂化汽油) 和 3 种最终产品可以定义模型的变量. 令

$$x_{ij} = \text{用来混合成最终产品 } j \text{ 的输入流 } i \text{ 的数量 (桶/天)}, \quad i=1,2; \quad j=1,2,3$$

根据这个定义, 我们有

$$\text{普通汽油的日产量} = (x_{11} + x_{21}) \text{ 桶/天}$$
$$\text{优质汽油的日产量} = (x_{12} + x_{22}) \text{ 桶/天}$$
$$\text{超优汽油的日产量} = (x_{13} + x_{23}) \text{ 桶/天}$$

$$\begin{aligned}\begin{pmatrix}\text{混合装置}\\\text{的日输出}\end{pmatrix} &= \begin{pmatrix}\text{普通汽油}\\\text{的日产量}\end{pmatrix} + \begin{pmatrix}\text{优质汽油}\\\text{的日产量}\end{pmatrix} + \begin{pmatrix}\text{超优汽油}\\\text{的日产量}\end{pmatrix}\\ &= [(x_{11} + x_{21}) + (x_{12} + x_{22}) + (x_{13} + x_{23})] \text{桶/天}\end{aligned}$$

$$\text{送入混合装置的日存料量} = (x_{11} + x_{12} + x_{13}) \text{ 桶/天}$$
$$\text{裂化装置送入混合装置的日送入量} = (x_{21} + x_{22} + x_{23}) \text{ 桶/天}$$
$$\text{送入裂化装置的日存料量} = 2(x_{21} + x_{22} + x_{23}) \text{ 桶/天}$$

$$\text{精炼厂的日原油使用量} = [5(x_{11} + x_{12} + x_{13}) + 10(x_{21} + x_{22} + x_{23})] \text{ 桶/天}$$

模型的目标是极大化由 3 种等级汽油产生的总利润. 由上面给出的定义得

$$\max\ z = 6.70(x_{11} + x_{21}) + 7.20(x_{12} + x_{22}) + 8.10(x_{13} + x_{23})$$

建立问题的约束如下:

(1) 原油日提供量不超过 1 500 000 桶/天:

$$5(x_{11} + x_{12} + x_{13}) + 10(x_{21} + x_{22} + x_{23}) \leqslant 1\,500\,000$$

(2) 裂化装置的输入能力不超过 200 000 桶/天:

$$2(x_{21} + x_{22} + x_{23}) \leqslant 200\,000$$

(3) 普通汽油的日需求不超过 50 000 桶:

$$x_{11} + x_{21} \leqslant 50\,000$$

(4) 优质汽油的日需求不超过 30 000 桶:

$$x_{12} + x_{22} \leqslant 30\,000$$

(5) 超优汽油的日需求不超过 40 000 桶:

$$x_{13} + x_{23} \leqslant 40\,000$$

(6) 普通汽油的辛烷值 (ON) 不低于 87:

汽油产品的辛烷值是混合过程中输入流的辛烷值的加权平均, 可计算如下:

$$\begin{aligned}&\text{普通汽油的平均 ON} =\\ &\frac{\text{存料的 ON} \times \text{存料桶数/天} + \text{裂化汽油的 ON} \times \text{裂化汽油的桶数/天}}{\text{普通汽油的总桶数/天}}\\ &= \frac{82x_{11} + 98x_{21}}{x_{11} + x_{21}}\end{aligned}$$

因此, 普通汽油辛烷数的约束是

$$\frac{82x_{11}+98x_{21}}{x_{11}+x_{21}} \geqslant 87$$

将约束线性化后得到

$$82x_{11}+98x_{21} \geqslant 87(x_{11}+x_{21})$$

(7) 优质汽油的辛烷值 (ON) 不低于 89:

$$\frac{82x_{12}+98x_{22}}{x_{12}+x_{22}} \geqslant 89$$

线性化得到

$$82x_{12}+98x_{22} \geqslant 89(x_{12}+x_{22})$$

(8) 超优汽油的辛烷值 (ON) 不低于 92:

$$\frac{82x_{13}+98x_{23}}{x_{13}+x_{23}} \geqslant 92$$

或

$$82x_{13}+98x_{23} \geqslant 92(x_{13}+x_{23})$$

因此, 完整的模型概括如下:

$$\text{max;}\overline{\text{mize}}\ z = 6.70(x_{11}+x_{21})+7.20(x_{12}+x_{22})+8.10\ \ (x_{13}+x_{23})$$

$$\begin{aligned}
\text{s.t.}\quad 5(x_{11}+x_{12}+x_{13})+10(x_{21}+x_{22}+x_{23}) &\leqslant 1\,500\,000 \\
2(x_{21}+x_{22}+x_{23}) &\leqslant 200\,000 \\
x_{11}+x_{21} &\leqslant 50\,000 \\
x_{12}+x_{22} &\leqslant 30\,000 \\
x_{13}+x_{23} &\leqslant 40\,000 \\
82x_{11}+98x_{21} &\geqslant 87(x_{11}+x_{21}) \\
82x_{12}+98x_{22} &\geqslant 89(x_{12}+x_{22}) \\
82x_{13}+98x_{23} &\geqslant 92(x_{13}+x_{23}) \\
x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23} &\geqslant 0
\end{aligned}$$

可以简化最后 3 个约束, 生成常数右端项.

解 最优解 (用文件 amplEx2.3-7.txt 求解) 为 $z = 875\ 000$, $x_{11} = 34\ 375$, $x_{21} = 15\ 625$, $x_{12} = 16\ 875$, $x_{22} = 13\ 125$, $x_{13} = 15\ 000$, $x_{23} = 25\ 000$. 将结果转述成

日利润 = 875 000 美元

普通汽油的日产量 $= x_{11} + x_{21} = 34\,375 + 15\,625 = 50\,000$ 桶/天

优质汽油的日产量 $= x_{12} + x_{22} = 16\,875 + 13\,125 = 30\,000$ 桶/天

超优汽油的日产量 $= x_{13} + x_{23} = 15\,000 + 25\,000 = 40\,000$ 桶/天

这个结果表明, 所有等级的需求能够满足.

习题 2.3E

1. Hi-V 使用新鲜的草莓、葡萄和苹果生产 3 种类型的罐装果汁饮料 A, B, C. 原料的日供应上限是 200 吨草莓、100 吨葡萄和 150 吨苹果. 每吨草莓、葡萄和苹果的成本分别为 200 美元、100 美元和 90 美元. 每吨原料可分别做成 1 500 磅草莓汁、1 200 磅葡萄汁和 1 000 磅苹果汁. 饮料 A 是 1:1 的草莓汁和苹果汁的混合. 饮料 B 是 1:1:2 的草莓汁、葡萄汁和苹果汁的混合. 饮料 C 是 2:3 的葡萄汁和苹果汁的混合. 所有的饮料按 16 盎斯 (1 磅)/罐被罐装起来. 对于饮料 A, B, C, 每一罐的价格分别是 1.15 美元、1.25 美元和 1.20 美元. 试确定这 3 种饮料的最优生产混合.

***2.** 一间五金店包装的杂工包由螺丝钉、螺栓、螺母和垫圈构成. 螺丝钉装在 100 磅盒中, 每盒的成本是 110 美元; 螺栓装在 100 磅盒中, 每盒的成本 150 美元; 螺母装在 80 磅盒中, 每盒的成本 70 美元; 垫圈装在 30 磅盒中, 每盒的成本 20 美元. 杂工包至少重 1 磅, 按重量至少要包括 10% 的螺丝钉与 25% 的螺栓, 最多有 15% 的螺母与 10% 的垫圈. 为了平衡包装袋, 螺栓的数量不能够超过螺母的数量或垫圈的数量. 螺栓的重量是螺母重量的 10 倍, 是垫圈重量的 50 倍. 试确定包装袋的最优混合.

3. All-Natural 公司从 4 种配料 (燕麦片、葡萄干、椰丝和长杏仁) 生产 3 种早餐麦片食品 A, B, C, 配料的日可用量分别是 5 吨、2 吨、1 吨和 1 吨. 每吨的相应成本分别为 100 美元、120 美元、110 美元和 200 美元. 麦片 A 中燕麦片、葡萄干和杏仁的混合比例是 50:5:2. 麦片 B 中燕麦片、椰子和杏仁混合比例是 60:2:3. 麦片 C 中燕麦片、葡萄干、椰子和杏仁的混合比例是 60:3:4:2. 食品被制成 5 磅大盒包装. All-Natural 销售 A, B, C 这 3 种麦片的价格分别是每盒 2 美元、2.50 美元和 3.00 美元. 对麦片 A, B, C 的最小日需求量分别是 500 盒、600 盒和 500 盒. 试确定这 3 种麦片最优生产混合以及相应的配料数量.

4. 某精炼厂生产 F1 和 F2 两个等级的喷气式飞机燃料, 由 4 种类型的汽油 A, B, C, D 混合而成. 燃料 F1 使用 4 种类型的汽油 A, B, C, D 的比例是 1:1:2:4, 燃料 F2 使用的比例是 2:2:1:3. 4 种汽油 A, B, C, D 的最大供应量分别是 1 000, 1 200, 900, 1 500 桶/天, 每桶成本分别是 120 美元、90 美元、100 美元和 150 美元. 燃料 F1 和 F2 的售价是每桶 200 美元和 250 美元. 燃料 F1 和 F2 的最小需求量是 200 和 400 桶/天. 试确定 F1 和 F2 的最优生产混合.

***5.** 一家汽油公司蒸馏 A 和 B 两种原油, 生产普通汽油、优质汽油以及航空燃油. 对原油的日供应量以及最终产品最小需求量均有限制. 如果生产不能满足需求, 不足的部分采取外购补充, 并会受到经济惩罚. 剩余产品不能立即销售, 会带来相应的储存成本. 下表提供了这种情形的数据:

原油	每桶产生的部分			价格/桶(美元)	桶/天
	普通	优质	航空		
原油 A	0.20	0.10	0.25	30	2 500
原油 B	0.25	0.30	0.10	40	3 000
需求 (桶/天)	500	700	400		
收入 (美元/桶)	50	70	120		
剩余产品储存费用 (美元/桶)	2	3	4		
未满足需求的惩罚 (美元/桶)	10	15	20		

试为精炼厂确定最优的产品混合.

6. 在第 5 题的精炼厂情形中, 假设蒸馏装置实际上生产中间产品石脑油和轻油. 1 桶原油 A 生产 0.35 桶石脑油和 0.6 桶轻油, 而 1 桶原油 B 生产 0.45 桶石脑油和 0.5 桶轻油. 石脑油和轻油被混合后生产 3 种最终的汽油产品: 普通汽油的混合比例是 2:1 (石脑油比轻油), 优质汽油的混合比例是 1:1, 航空燃油的混合比例是 1:2. 试确定最优的产品混合.

7. 夏威夷糖业公司通过甘蔗汁来生产红糖、精加工 (白) 糖、糖粉, 以及糖蜜. 公司每星期购买甘蔗汁 4 000 吨, 按照合同, 每种糖类制品每周至少生产 25 吨. 生产过程是从甘蔗汁中制造红糖和糖蜜开始. 1 吨甘蔗汁可生产 0.3 吨红糖和 0.1 吨糖蜜. 白糖是由红糖加工生产的. 用 1 吨红糖可生产 0.8 吨的白糖. 糖粉由白糖通过一个特殊的研磨过程生产而成, 其转化率为 95%(1 吨白糖生产 0.95 吨糖粉). 对于红糖、白糖、糖粉和糖蜜每吨的利润分别是 150 美元、200 美元、230 美元和 35 美元. 建立该问题的线性规划表达式, 并且确定每周的生产进度.

8. Shale 石油精炼厂混合 A 和 B 两种石油原料, 生产 I 和 II 两种高辛烷值的汽油产品, 生产原料 A 和 B 的最大生产率分别是 450 和 700 桶/小时. 相应的辛烷值分别是 98 和 89, 而且蒸汽压力分别是 10 和 8 磅/英寸^2. 汽油 I 和汽油 II 的辛烷值分别要至少达到 91 和 93. 与两产品相关的蒸汽压力应该不超过 12 磅/英寸^2. 产品 I 和 II 的每桶利润分别为 7 美元和 10 美元. 对于产品 I 和 II 确定最优的生产率, 以及原料 A 和原料 B 的混合比例. (**提示**: 蒸汽压力与辛烷值一样, 是混合原料蒸汽压力的加权平均值).

9. 一铸造厂熔炼钢、铝、铸铁废料生产 I 和 II 两种类型的金属锭铁, 而且对铝、石墨、硅成分有特殊的限制. 铝块和硅块用于在熔炼过程中达到所需的规格要求. 下表设定了本问题的规格:

输入项	成分 (%)			费用/吨(美元)	可用量(吨/天)
	铝	石墨	硅		
钢废料	10	5	4	100	1 000
铝废料	95	1	2	150	500
铸铁废料	0	15	8	75	2 500
铝块	100	0	0	900	任意量
硅块	0	0	100	380	任意量

成分	锭铁 I		锭铁 II	
	最小值	最大值	最小值	最大值
铝	8.1%	10.8%	6.2%	8.9%
石墨	1.5%	3.0%	4.1%	∞
硅	2.5%	∞	2.8%	4.1%
需求 (吨/天)	130		250	

为铸造厂熔炼废料确定最优的输入混合.

10. A 与 B 两类合金, 由 Ⅰ, Ⅱ, Ⅲ, Ⅳ 这 4 种金属按照下列规格制成:

合金	规格	售价 (美元)
A	至多 80% 的 Ⅰ	200
	至多 30% 的 Ⅱ	
	至少 50% 的 Ⅳ	
B	40%～60% 的 Ⅱ	300
	至少 30% 的 Ⅲ	
	至少 70% 的 Ⅳ	

这 4 种金属按照下列数据, 从 3 种矿石中提取:

矿石	最大量 (吨)	成分 (%)					价格/吨 (美元)
		Ⅰ	Ⅱ	Ⅲ	Ⅳ	其他	
1	1 000	20	10	30	30	10	30
2	2 000	10	20	30	30	10	40
3	3 000	5	5	70	20	0	50

每种类型的合金应该生产多少? (提示: 令 x_{jk} 是矿石 j 分配给合金 k 的吨数, 并定义 w_k 为所生产的合金 k 的吨数.)

2.3.6 人力规划

正如例 2.3-6 所示范的那样, 可以通过雇用和解雇的办法来满足不同时期对于劳动力的需求. 有些情况下, 需求浮动的影响可以通过调整一个工作班组的上下班时间来“吸收”掉. 例如, 可以不遵循传统分三班、每班工作 8 小时 (上班开始时间分别是早上 8:00、下午 4:00 和晚上 12:00) 的工作方法, 我们可以使用交错的 8 小时一班的方法, 制定每一班工作的开始时间以适应需求的增加或减少.

重新确定一班的开始以适应需求浮动的这一思想也可以扩充到其他的作业环境. 例 2.3-8 的问题是, 在满足高峰时间和低峰时间的运输需求的情况下, 求所需的最少公交车数.

实际应用——Qantas 航空公司电话销售人力规划

澳大利亚 Qantas 航空公司订票处的运营时间是从早 7:00 到晚 10:00, 共有 6 个班组, 每班组在一天的不同时间开始上班. Qantas 为了给客户提供方便的服务,

使用线性规划 (嵌入排队分析) 高效地为订票处配置员工. 在 20 世纪 70 年代末完成的这项研究, 每年为公司节省 200 000 澳元. 第 24 章的案例 15 提供了研究的详细资料.

例 2.3-8 (公交车调度安排)

Progress 市正在研究关于引进公交系统以减轻城市内自驾车引起的烟尘污染的可行性, 这项研究是寻求能满足运输所需的最少公交车数. 在收集了必要的信息之后, 市政工程师注意到, 所需的最少公交车数随一天中的时间不同而变化, 而且所需要的最少公交车数在若干连续的 4 小时间隔内可以被近似为一个常数. 图 2.8 概述了工程师的发现. 为了完成所需的日常维护, 每辆公交车一天只能连续运行 8 小时.

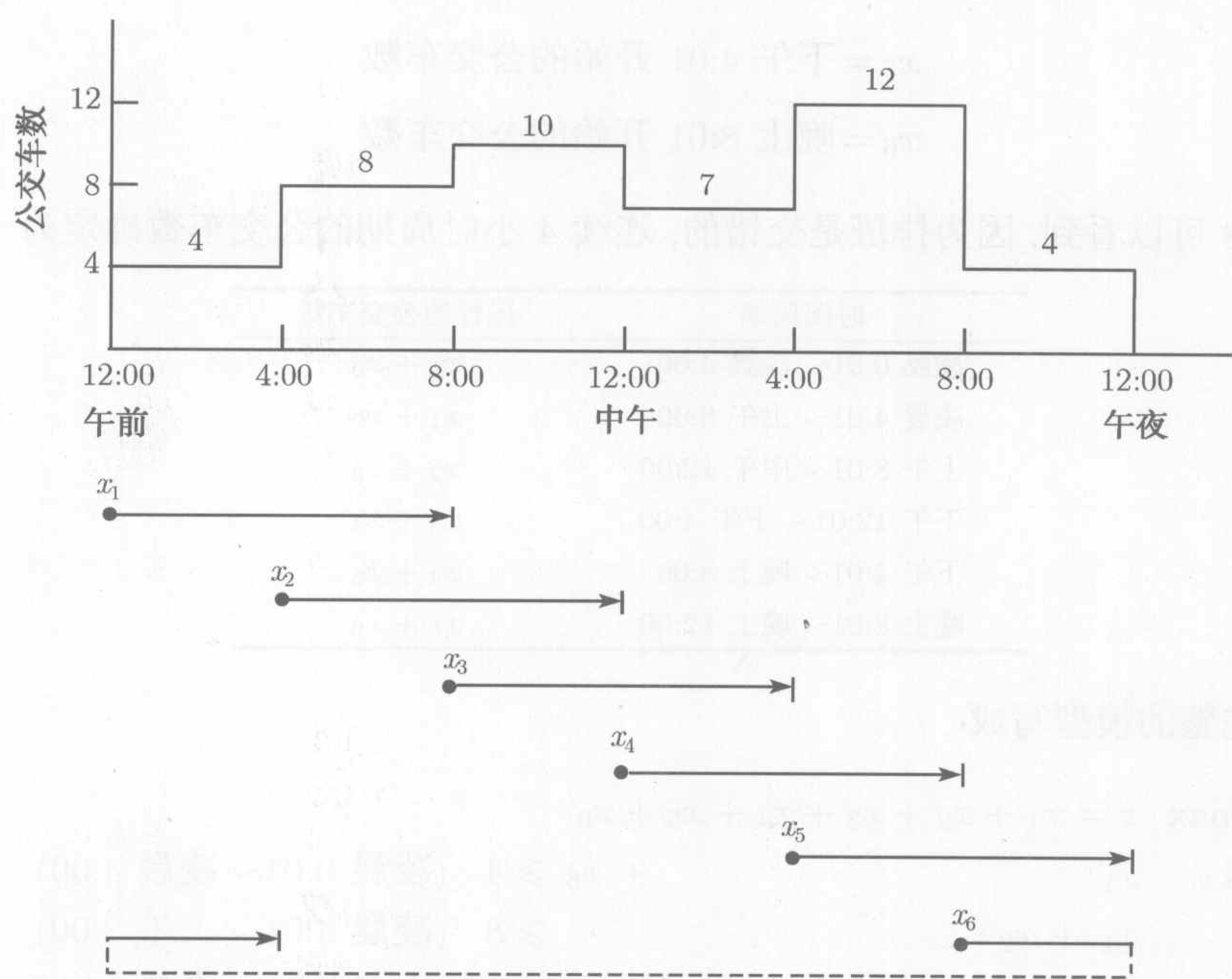


图 2.8 公交车数可认为是一天时间的函数

数学模型 确定每一班运行公交车的数量 (变量), 以满足最小需求 (约束), 使所运行的公交车总数最少 (目标).

你可能已注意到, 变量的定义是不明确的. 我们知道, 每辆公交车将连续运行 8 小时, 但我们并不知道一个班开始的时间. 如果遵循正常的三班时间表 (上午 8:01~下午 4:00, 下午 4:01 ~ 午夜 12:00, 凌晨 0:01 ~ 上午 8:00), 并假定 x_1, x_2 和 x_3 是第一班、第二班和第三班公交车开始时的数量, 我们可以从图 2.8 看到, $x_1 \geqslant 10$,

$x_2 \geqslant 12, x_3 \geqslant 8$. 相应的日公交车的最少数量是 $x_1 + x_2 + x_3 = 10 + 12 + 8 = 30$.

上述结果是可接受的, 仅当排班必须与正常的三班时间表相符. 然而, 最好让最优化过程来选择每一班的“最好”的开始时间. 完成这一目标的合理方法是允许每 4 小时开始一班. 图 2.8 的底部解释了这一概念, 重叠的 8 小时班开始时刻是凌晨 0:01、凌晨 4:01、上午 8:01、下午 12:01、下午 4:01 和晚上 8:01. 因此, 变量可以定义为

$x_1 =$ 凌晨 0:01 开始的公交车数

$x_2 =$ 凌晨 4:01 开始的公交车数

$x_3 =$ 上午 8:01 开始的公交车数

$x_4 =$ 下午 12:01 开始的公交车数

$x_5 =$ 下午 4:01 开始的公交车数

$x_6 =$ 晚上 8:01 开始的公交车数

从图 2.8 可以看到, 因为排班是交错的, 连续 4 小时周期的公交车数确定为

时间周期	运行的公交车数
凌晨 0:01～ 凌晨 4:00	$x_1 + x_6$
凌晨 4:01～ 上午 8:00	$x_1 + x_2$
上午 8:01～ 中午 12:00	$x_2 + x_3$
下午 12:01～ 下午 4:00	$x_3 + x_4$
下午 4:01～ 晚上 8:00	$x_4 + x_5$
晚上 8:01～ 晚上 12:00	$x_5 + x_6$

因此, 完整的模型写成

$$
\begin{aligned}
\max\quad & z = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \\
\text{s.t.}\quad & x_1 + x_6 \geqslant 4 \quad (\text{凌晨 } 0{:}01\sim \text{凌晨 } 4{:}00) \\
& x_1 + x_2 \geqslant 8 \quad (\text{凌晨 } 4{:}01\sim \text{上午 } 8{:}00) \\
& x_2 + x_3 \geqslant 10 \quad (\text{上午 } 8{:}01\sim \text{中午 } 12{:}00) \\
& x_3 + x_4 \geqslant 7 \quad (\text{下午 } 12{:}01\sim \text{下午 } 4{:}00) \\
& x_4 + x_5 \geqslant 12 \quad (\text{下午 } 4{:}01\sim \text{晚上 } 8{:}00) \\
& x_5 + x_6 \geqslant 4 \quad (\text{晚上 } 8{:}01\sim \text{晚上 } 12{:}00) \\
& x_j \geqslant 0,\ j = 1, 2, \cdots, 6
\end{aligned}
$$

解　最优解要求使用 26 辆公交车就可以满足需求, 其中 $x_1 = 4$, 开始时间在凌晨 0:01; $x_2 = 10$, 从凌晨 4:01 开始; $x_4 = 8$, 从下午 12:01 开始; 以及 $x_5 = 4$, 从下午 4:01 开始.

习题 2.3F

*1. 在公交车调度安排的例子中, 假定公交车可以实行 8 小时/班或 12 小时/班的工作制度. 如果一辆公交车运行 12 小时, 必须付给司机的加班工资是正常班工资的 150%. 你推荐使用 12 小时/班的工作制度吗?

2. 一家医院雇用志愿者作为接待处的工作人员, 接待时间是从早上 8:00 到晚上 10:00. 每名志愿者连续工作 3 小时, 只有在晚上 8:00 开始工作的人员除外, 他们只工作 2 小时. 对于志愿者的最小需求可以近似成以 2 小时作为间隔的阶梯函数: 从早上 8:00 开始, 相应的需求人数分别是 4, 6, 8, 6, 4, 6, 8. 因为大多数志愿者是退休人员, 他们愿意在一天的任何时间 (早上 8:00~ 晚上 10:00) 提供服务. 然而, 由于大多数慈善团体竞争他们的服务, 所需的数目必须尽可能少. 为志愿者的开始时间确定最优安排.

3. 在第 2 题中, 考虑到午饭和晚饭, 假定没有志愿者愿意在中午 12:00 或晚上 6:00 开始工作, 确定最优安排.

4. 在 LTL 货车运输公司, 装卸点有一些临时工人, 最繁忙的时候临时雇用他们装卸货物. 在内布拉斯加州的 Omaha 装卸点, 在一周 7 天内 (星期一开始) 对临时工最小需求是 20, 14, 10, 15, 18, 10, 12 名工人. 与每名工人签订连续 5 天工作的合同. 为公司制定每周雇用临时工人的最优方案.

*5. 大多数大学在校学生与所在院系签订合同, 为学校做些如接听电话和打字等杂活. 对这类服务的需求在整个工作时间内 (上午 8:00~ 下午 5:00) 是波动的. 在工业工程系, 所需学生的最少数量在上午 8:00~ 上午 10:00 之间是 2 名, 在上午 10:01~ 上午 11:00 之间是 3 名, 在上午 11:01~ 下午 1:00 之间是 4 名, 在下午 1:01~ 下午 5:00 之间是 3 名. 分配每位学生连续工作 3 小时 (在下午 3:01 开始工作的, 他们连续工作 2 小时; 在下午 4:01 开始工作的, 他们工作 1 小时). 由于学生们的时间很灵活, 他们可以在工作日随时开始上班, 午饭时间 (中午 12:00) 除外, 学生不打算在此时开始工作. 求工业工程系应该雇用的最少学生数, 并指定他们的工作开始时间.

6. 一家大的百货公司一周营业 7 天. 经理估计能够提供及时服务所需售货员的最少数量是周一 12 名, 周二 18 名, 周三 20 名, 周四 28 名, 周五 32 名, 周六和周日各 40 名. 每名售货员一周工作 5 天, 连续休息 2 天, 并在一周内相互错开. 例如, 如果 10 名售货员在周一开始工作, 有 2 名可以在周二和周三休息, 5 名可在周三和周四休息, 3 名可在周六和周日休息. 应雇用多少售货员, 应该分配他们在哪些天休息?

2.3.7 其他应用

前面的各节介绍了线性规划应用的 6 个具有代表性的领域. 事实上在众多领域内线性规划都能得到很好的应用. 本节最后所列的习题显示其中一些应用领域, 其应用范围从农业到军事各个方面. 本节还介绍一个有趣的应用, 即如何按照顾客指定的尺寸来切割标准的原料纸卷问题.

例 2.3-9 (下料问题或原料切割问题)

太平洋纸业公司生产纸卷, 每卷的标准宽度为 20 英尺. 特殊顾客对不同宽度的要求可以通过切割标准卷来生产. 典型的订单要求 (它可以每日不同) 如下表:

要求	期望宽度 (英尺)	要求卷数
1	5	150
2	7	200
3	9	300

在实际中, 可以通过把切刀设置成所需的宽度以满足订单要求. 通常, 对于一个给定的订单要求, 可以为标准纸卷设置许多切割方案. 对于 20 英尺宽的纸卷, 图 2.9 演示了 3 种可行的切割设置. 虽然还有其他可行的设置, 但我们仅限于对图 2.9 中设置 1、设置 2 和设置 3 这 3 种情形进行讨论. 我们可以用多种方式组合给定的设置, 以满足宽度为 5 英尺、7 英尺和 9 英尺的要求. 下面是可行组合的例子:

组合 1: 用设置 1 切割 300 个 (标准) 纸卷, 用设置 2 切割 75 个纸卷.

组合 2: 用设置 1 切割 200 个纸卷, 用设置 3 切割 100 个纸卷.

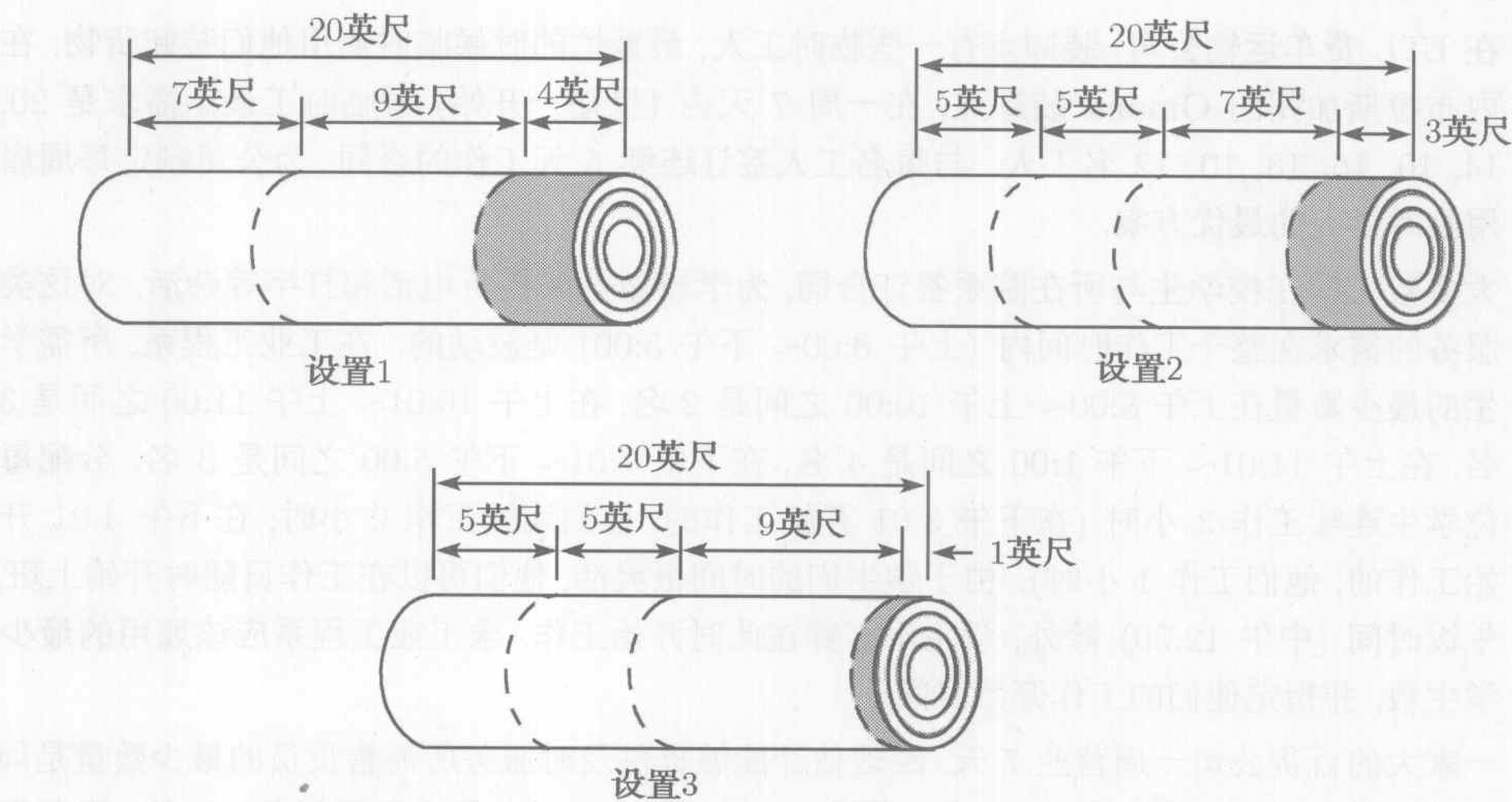


图 2.9 对于切刀设置 1, 2, 3 的下料损失 (阴影部分)

哪一种组合更好呢? 可以考虑每种组合产生的"浪费"来回答这个问题. 图 2.9 中的阴影部分表示不能满足订单需求的宽度而剩余的纸卷. 这些剩余纸卷被称为下料损失(trim loss). 可以通过计算它的下料损失来估计每种组合的"优度". 然而, 因为剩余纸卷也许有不同的宽度, 我们应该根据下料损失的**面积**而不是剩余纸卷的**数量**来估计. 假定标准卷的长度是 L 英尺, 我们可以计算下料损失的面积如下:

$$\text{组合 1: } 300(4 \times L) + 75(3 \times L) = 1\,425L \text{ 英尺}^2$$

$$\text{组合 2: } 200(4 \times L) + 100(1 \times L) = 900L \text{ 英尺}^2$$

这些面积仅计算图 2.9 中的阴影部分. 任何 5 英尺宽纸卷、7 英尺宽纸卷和 9 英尺宽卷的剩余产品也必须考虑在下料损失面积的计算中. 在组合 1 中, 设置 1 剩

余额外的 7 英尺宽纸卷是 $300-200=100$ 卷, 设置 2 剩余额外 7 英尺宽纸卷 75 卷. 因此, 额外的浪费面积是 $175(7\times L)=1\,225L$ 英尺^2. 组合 2 不产生 7 英尺宽纸卷和 9 英尺宽纸卷的剩余, 但设置 3 确实产生额外的 5 英尺宽纸卷 $200-150=50$ 卷, 也就是附加上 $50(5\times L)=250L$ 英尺^2 的浪费面积. 其结果, 我们有

$$\text{组合 1 总的下料损失面积} = 1\,425L + 1\,225L = 2\,650L \text{ 英尺}^2$$

$$\text{组合 2 总的下料损失面积} = 900L + 250L = 1\,150L \text{ 英尺}^2$$

因此组合 2 更好一点, 因为它产生较少的下料损失面积.

数学模型 问题可以用文字概括为: *求满足所需订单要求*(约束) *的切刀的设置组合*(变量), *使下料损失面积为最小*(目标).

上面的变量一定要按照切割操作员能使用的方式给出定义. 特别地, 将变量定义为*依照给定的切刀设置而切割出的标准纸卷数量*. 这个定义需要确定所有可能的切刀设置, 其概述在下表中 (设置 1、设置 2 和设置 3 是图 2.9 给出的情况). 你应该确信, 设置 4、设置 5 和设置 6 是有效的, 并且这包括了所有"可能的"设置. 记住, 一种可能的设置产生的下料损失纸卷的宽度不能大于等于 5 英尺.

所需宽度 (英尺)	刀的设置						最小纸卷数
	1	2	3	4	5	6	
5	0	2	2	4	1	0	150
7	1	1	0	0	2	0	200
9	1	0	1	0	0	2	300
下料损失长度 (英尺)	4	3	1	0	1	2	

为了用数学方法表示该模型, 我们定义变量如下:

$$x_j = \text{按照设置 } j \text{ 而切割出的标准纸卷的数量}, \; j=1,2,\cdots,6$$

模型的约束即要满足纸卷的需求.

$$\text{产生 5 英尺宽纸卷数量} = 2x_2 + 2x_3 + 4x_4 + x_5 \geqslant 150$$

$$\text{产生 7 英尺宽纸卷数量} = x_1 + x_2 + 2x_5 \geqslant 200$$

$$\text{产生 9 英尺宽纸卷数量} = x_1 + x_2 + x_3 + 2x_5 + 2x_6 \geqslant 300$$

为构造目标函数, 我们注意到, 总的下料损失面积是已使用的标准卷总面积与表示所有订单的总面积两者之差. 因此,

$$\text{标准纸卷总面积} = 20L(x_1 + x_2 + x_3 + x_4 + x_5 + x_6)$$

$$\text{订单总面积} = L(150\times 5 + 200\times 7 + 300\times 9) = 4\,850L$$

目标函数则变为

$$\min \; z = 20L(x_1 + x_2 + x_3 + x_4 + x_5 + x_6) - 4\,850L$$

由于标准纸卷宽度长 L 是一常数, 目标函数等价地简化为

$$\min\ z = x_1 + x_2 + x_3 + x_4 + x_5 + x_6$$

因此, 模型可写成

$$\begin{aligned}
\min\ & z = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \\
\text{s.t.}\ & 2x_2 + 2x_3 + 4x_4 + x_5 \geqslant 150 \quad (5\ \text{英尺宽纸卷}) \\
& x_1 + x_2 + 2x_5 \geqslant 200 \quad (7\ \text{英尺宽纸卷}) \\
& x_1 + x_2 + x_3 + 2x_5 + 2x_6 \geqslant 300 \quad (9\ \text{英尺宽纸卷}) \\
& x_j \geqslant 0,\ j = 1, 2, \cdots, 6
\end{aligned}$$

解 最优解要求, 按照设置 4 切割 12.5 个标准纸卷, 按照设置 5 切 100 个标准纸卷, 以及按照设置 6 切 150 个标准纸卷. 这个解并不能实现, 因为变量 x_4 不是整数. 可以用整数算法求解此问题 (见第 8 章), 或者对 x_4 采用四舍五入的原则, 切割 13 卷.

评注 这里介绍的下料模型假定, 所有可行的切刀设置是预先给定的. 这项工作对于大型问题可能是困难的, 并且可以实施的可行组合有可能被错过. 这个问题可以由一个嵌入在线性规划模型中的整数规划予以完善, 用来根据需求来生成所有可能的切刀设置, 直到找出最优解为止. 这种算法有时称为**列生成** (column generation), 附录 E(见下册) 的综合问题 13-3 中给出了详细的论述. 此方法由 (稍微高级的) 线性规划理论来确立, 并足以证明线性规划理论在实践中的重要性.

习题 2.3G

***1.** 考虑例 2.3-9 的下料模型.

(a) 如果用设置 1 切割 200 个纸卷, 用设置 3 切割 100 个纸卷, 计算相应的下料损失面积.

(b) 假设我们只有 15 英尺宽的标准纸卷. 为 5 英尺宽纸卷、7 英尺宽纸卷和 9 英尺宽纸卷生成所有可能的切刀设置, 并计算每英尺长纸卷相应的下料损失.

(c) 在原来的模型中, 如果 7 英尺宽纸卷的需求减少 80 卷, 必须满足 3 种类型纸卷需求的标准 20 英尺宽纸卷的最小数量是多少?

(d) 在原来的模型中, 如果对 9 英尺宽纸卷的需求变到 400 卷, 必须满足新的需求的 20 英尺宽标准纸卷应该增加多少?

2. 货架的空间分配. 一食品商店必须决定 5 种早餐麦片在货架上的空间分配. 这 5 种早餐麦片的最大日需求量分别是 100 箱、85 箱、140 箱、80 箱和 90 箱. 每种箱子分别占货架面积 16 平方英寸、24 平方英寸、18 平方英寸、22 平方英寸和 20 平方英寸. 总的可用货架面积是 5 000 平方英寸. 每种麦片单位利润分别是 1.10 美元、1.30 美元、1.08 美元、1.25 美元和 1.20 美元. 为 5 种麦片确定最优的面积分配.

3. 议题投票. 在阿肯色州的某县, 4 个竞选议题需要投票: 建设新公路、加大枪支的控制、增加农场补助金和提高汽油税. 该县总共有 100 000 位城市选民、250 000 位市郊选民和

50 000 位农村选民, 他们对这些竞选议题的支持度和反对程度是多样的. 例如, 农村选民反对枪支控制和提高汽油税, 但支持道路建设和农场补助金. 该县正计划一项电视广告活动, 广告的预算是 100 000 美元, 而每次广告的成本是 1 500 美元. 下表概括了单项广告的影响, 其依据是对于不同议题的赞成票数和反对票数:

议题	每项广告赞成 (+) 和反对 (−) 票数的期望值		
	城市	市郊	农村
建设新公路	−30 000	+60 000	+30 000
枪支控制	+80 000	+30 000	−45 000
烟雾控制	+40 000	+10 000	0
汽油税	+90 000	0	−25 000

如果争取到至少 51% 的赞成票, 一项议题将被表决通过. 哪项议题将被投票通过, 应该有多少广告被分配到这些议题中?

4. **装配线的平衡.** 某产品由 3 个不同的零件装配而成. 零件是由两个部门以不同的生产率制造, 其值由下表给出:

部门	生产能力 (小时/周)	生产率 (件/小时)		
		零件 1	零件 2	零件 3
1	100	8	5	10
2	80	6	12	4

求每周内能够生产的最终产品的最大件数. (**提示:** 装配件数 = min{ 零件 1 的件数, 零件 2 的件数, 零件 3 的件数 }. max $z = \min\{x_1, x_2\}$ 等价于 max z s.t. $z \leqslant x_1, z \leqslant x_2$.)

5. **污染控制.** 有 3 种类型的煤 C1, C2, C3 被研磨成煤粉, 然后混合在一起, 其生产能力是每小时 50 吨, 以满足发电厂的燃料需要. 煤燃烧排出二氧化硫 ($\mu L/L$) 必须符合环保署 (EPA) 的规定, 至多是 2 000 $\mu L/L$. 下表概述本问题的数据:

	C1	C2	C3
硫 ($\mu L/L$)	2 500	1 500	1 600
粉碎机能力 (吨/小时)	30	30	30
每吨煤的成本 (美元)	30	35	33

确定煤的最优混合.

***6.** **交通灯控制.** (Stark and Nicholes, 1972) 在 3 条高速公路 H1, H2, H3 行驶的汽车在它们离开高速路进入收费口叉道之前必须停下来等待绿灯. 从 H1, H2, H3 驶来的每辆汽车的通行费分别是 3 美元、4 美元和 5 美元. 来自 H1, H2, H3 的汽车流量分别是每小时 500, 600, 400 辆. 红绿灯的转换周期不超过 2.2 分钟, 而且在任何一条公路上的绿灯持续时间必须至少是 25 秒. 黄灯持续时间是 10 秒. 收费口每小时最大能够处理 510 辆汽车. 假定没有汽车在黄灯下通行, 为 3 条公路确定最优的绿灯时间间隔, 使得收费口在每个红绿灯周期的收费最大.

7. **将经验数据拟合成直线 (回归).** 在为初学者开设的 10 周打字班上, 每位学生的平均打字速度 (每分钟字数) 是学习周数的函数, 如下表给出.

第 x 周	1	2	3	4	5	6	7	8	9	10
每分钟打字数 y	5	9	15	19	21	24	26	30	31	35

对给出的数据, 求最优拟合直线关系 $\hat{y}=ax+b$ 中的系数 a 和 b. (*提示*: 极小化理论值 $\hat{y}$ 与经验值 y 偏差的绝对值的和. $\min\ |x|$ 等价于 $\min\ z\ \text{ s.t. } z \leqslant x,\ z \geqslant -x$.)

8. *新高速公路平整土地*. (Stark and Nicholes, 1972) 阿肯色州高速公路部门正在计划在不平坦的地形上建设一条 10 英里长的公路, 地形剖面图如图 2.10 所示. 建筑地形的宽度大约是 50 码. 为了使问题简化, 地形剖面可以用阶梯函数来代替, 正如图形所描绘的那样. 使用重型机器, 将泥土从高地区域运抵并填充到低地区域. 还有两个土坑, 土坑 I 和土坑 II, 位于 10 英里路段的末端, 必要时, 可以从那搬运额外的泥土. 土坑 I 有 20 000 立方码容量的泥土, 土坑 II 有 15 000 立方码容量的泥土. 从土坑 I 和土坑 II 移动泥土的成本分别是每立方码 1.50 美元和 1.90 美元. 运输成本是每英里每立方码 0.15 美元, 使用重型机械装载载重卡车的成本是每立方码 0.20 美元. 这意味着, 1 立方码的泥土从土坑 I 搬运 1 英里所花的总成本是 $(1.5+0.20)+1\times 0.15=1.85$ 美元, 而 1 立方码的泥土从高处填向低处 1 英里的费用是 $0.20+1\times 0.15=0.35$ 美元. 为平整这 10 英里路段, 制定一项最小成本计划.

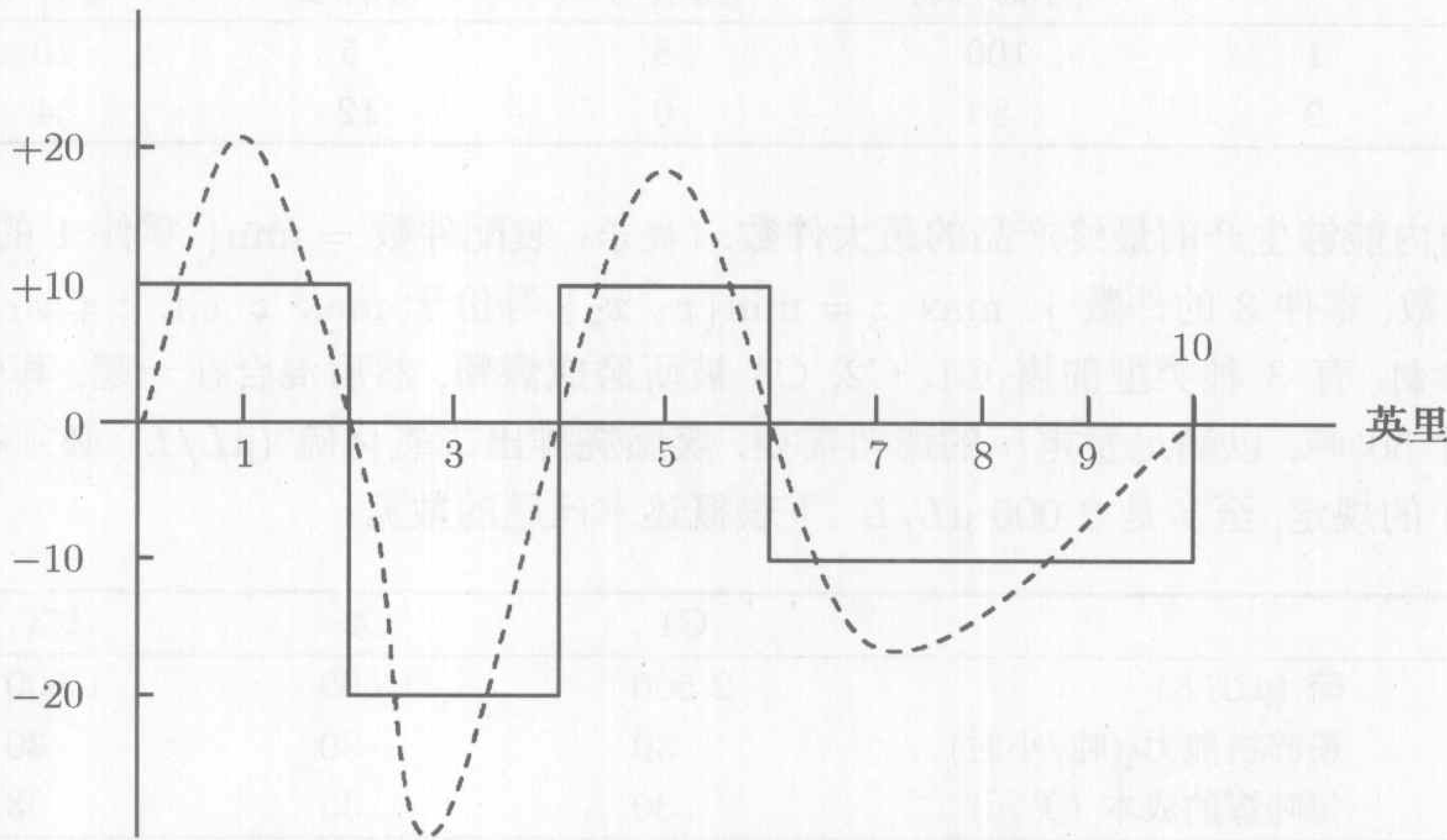


图 2.10　第 8 题的地形剖面

9. *军事作战计划*. (Shepard 等人, 1988) 红军 (R) 正在试图入侵由蓝军 (B) 防御的领地. 蓝军有 3 条防线和 200 个正规战斗单位, 并且还能够抽出 200 个预备单位. 红军计划进攻两条前线 (北线和南线); 蓝军设置 3 条东–西防线 (I, II, III). 防线 I 和 II 每条防线要至少阻止红军进攻 4 天以上, 并尽量延长总的战斗持续时间. 红军的前进时间由下列经验公式估计:

$$\text{战斗持续的天数} = a + b\left(\frac{\text{蓝军单位数}}{\text{红军单位数}}\right)$$

常数 a 和 b 是防线的函数, 并且北/南前线的情况由下表所示:

	a			b		
	Ⅰ	Ⅱ	Ⅲ	Ⅰ	Ⅱ	Ⅲ
北线	0.5	0.75	0.55	8.8	7.9	10.2
南线	1.1	1.3	1.5	10.5	8.1	9.2

蓝军的预备单位能够且只能用在防线 Ⅱ 上. 红军分配到 3 条防线的单位数由下表给出.

	红军进攻的单位数		
	防线 Ⅰ	防线 Ⅱ	防线 Ⅲ
北线	30	60	20
南线	30	40	20

蓝军应如何在北/南前线和 3 条防线上布署它的军队?

10. *水质管理*. (Stark and Nicholes, 1972) 4 个城市排出污水进入相同的河流. 城市 1 在上游, 下游依次是城市 2、城市 3、城市 4. 以河流方向估算, 城市之间的距离大约是 15 英里. 污水中所含污染物的衡量指标是 BOD(生化需氧量), 它描述的是稳定水中污物成分所需要的氧的重量. BOD 越高表明水质越差. 环保署设置了一个最大可允许的 BOD 含量, 用每加仑水中 BOD 的磅数来表示. 通过以下两种方式去除污水中污染物质: (1) 由空气中的氧气形成的自然分解作用; (2) 在污水到达河流之前, 建立排放点的污水处理工厂. 目标是确定 4 个污水处理厂每个厂最大经济效益, 约束是将 BOD 降低到可接受的水平. 工厂去污率的最大可能是 99%.

为说明这个过程的相关计算, 对工厂 1, 考虑下列定义:

Q_1 = 在通向城市 2 的河段 1~2 的 15 英里处的河水流量 (加仑/小时)

p_1 = BOD 排放率 (磅/小时)

x_1 = 工厂 1 的除污率 ($\leqslant 0.99$)

b_1 = 在河段 1~2 中, 所允许的最大 BOD 含量 (磅/加仑)

为满足河段 1~2 对 BOD 含量的要求, 必须有

$$p_1(1-x_1) \leqslant b_1Q_1$$

用类似的方法, 河段 2~3 的 BOD 含量约束有如下形式:

$$(1-r_{12})\begin{pmatrix}\text{在河段 1~2 中}\\ \text{BOD 的排放率}\end{pmatrix} + \begin{pmatrix}\text{在河段 2~3 中}\\ \text{BOD 的排放率}\end{pmatrix} \leqslant b_2Q_2$$

或者

$$(1-r_{12})p_1(1-x_1) + p_2(1-x_2) \leqslant b_2Q_2$$

系数 $r_{12}(<1)$ 表示在河段 1~2 中被分解掉的废物所占的百分数. 对于河段 2~3, 其约束是

$$(1-r_{23})[(1-r_{12})p_1(1-x_1) + p_2(1-x_2)] + p_3(1-x_3) \leqslant b_3Q_3$$

对于 4 个污水厂, 使用下列数据 (在所有 4 个河段中, 由分解去掉的废物所占的百分数是 6%), 确定最大的经济效果.

	河段 1~2 ($i=1$)	河段 2~3 ($i=2$)	河段 2~3 ($i=3$)	河段 3~4 ($i=4$)
Q_i (加仑/小时)	215 000	220 000	200 000	210 000
p_i (磅/小时)	500	3 000	6 000	1 000
b_i (BOD 磅数/加仑)	0.000 85	0.000 9	0.000 8	0.000 8
处理成本 (美元/去掉的 BOD 磅数)	0.20	0.25	0.15	0.18

11. *荷载结构*. (Stark and Nicholes, 1972) 如图 2.11 所示, 有两个吊车的高架起重机用于把混合后的混凝土吊运到某工地用来铸造混凝土障碍物. 装载混凝土的铲斗由吊车悬挂在正中央. 起重机两端的支撑点每个能够支撑的最大值是 25 千磅, 而吊车的每条钢丝绳具有 20 千磅的能力. 确定最大的负荷能力, W_1 和 W_2. (提示: 在平衡时, 在钢桁支架或吊车上任意一点的力矩总和为零.)

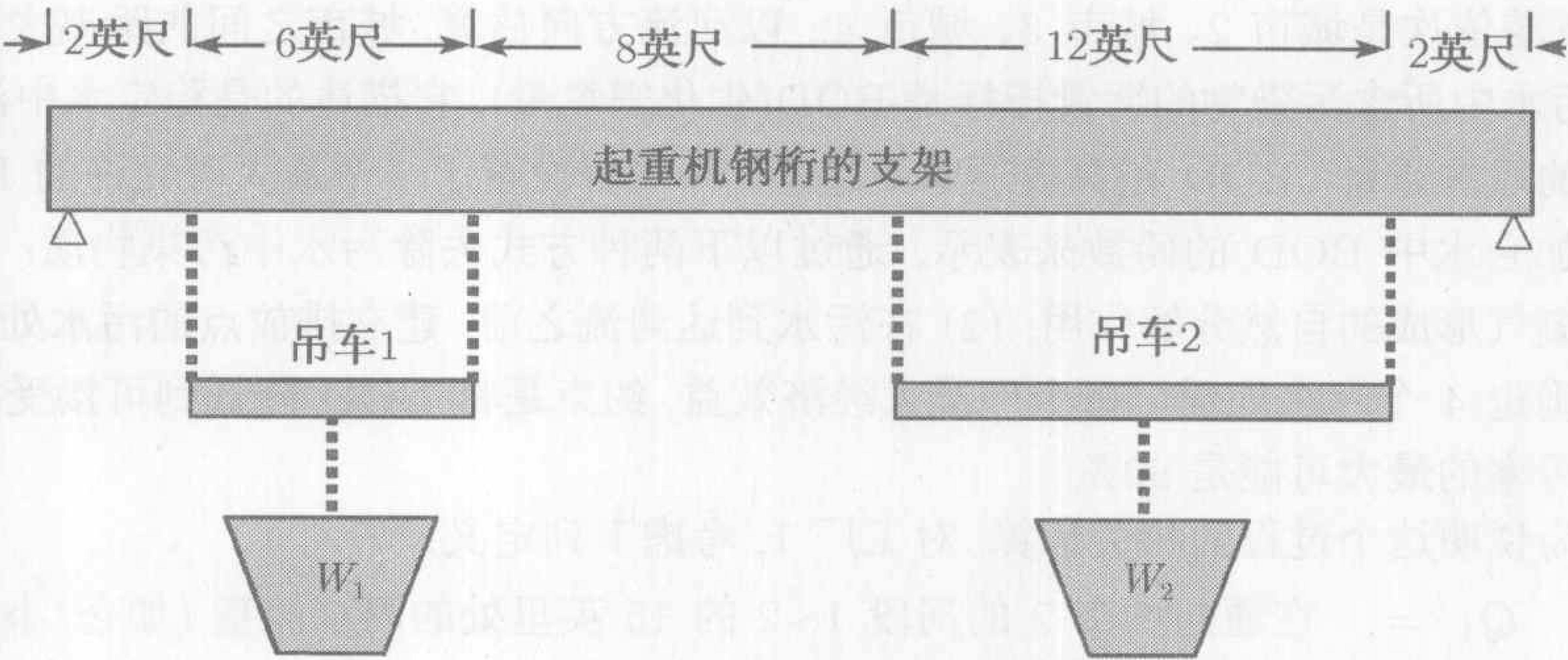


图 2.11　有两个吊车的高架起重机 (第 11 题)

12. *飞机分配*. 依照下面的数据, 考虑在 4 条线路上指派飞机:

飞机类型	载客能力 (乘客数)	飞机数	各条线路上的日飞行次数			
			1	2	3	4
1	50	5	3	2	2	1
2	30	8	4	3	3	2
3	20	10	5	5	4	2
每日顾客数			1 000	2 000	900	1 200

相应的费用, 包括由于满座而失去乘客的惩罚费用, 如下:

飞机类型	各条航线上每次飞行的运营成本 (美元)			
	1	2	3	4
1	1 000	1 100	1 200	1 500
2	800	900	1 000	1 000
3	600	800	800	900
失去一名乘客的罚款	40	50	45	70

确定飞机与航线之间的最优指派, 并确定相应的飞行次数.

2.4 借助于 Excel 规划求解和 AMPL 软件的计算机求解

在实际应用中, 典型的线性规划模型可能包括上千个变量和约束, 求解这些模型唯一可行的方法是使用计算机. 本节介绍两种不同的常用软件: Excel 规划求解和 AMPL 软件. 规划求解对于擅长电子表格的用户特别有吸引力. AMPL 软件是一种代数建模语言, 像任何其他的程序语言一样, 需要用户掌握专门的技能. 然而, AMPL, 以及其他一些类似的语言①, 在建模和运行大规模、复杂的线性规划模型方面提供了很大的灵活性. 尽管本节介绍的内容集中在线性规划上, 但 AMPL 和规划求解可以用于整数和非线性规划, 本书稍后将会给予介绍.

2.4.1 用 Excel 规划求解解线性规划问题

在 Excel 规划求解中, 电子表格是线性规划输入和输出的媒介. 图 2.12 显示了 Reddy Mikks 模型 (文件 solverRM1.xls) 的数据布局图. 该图形的顶部包含了 4 类信息: (1) 输入数据单元格 (阴影区域, B5:C9 和 F6:F9); (2) 描述我们寻求的变量和目标函数的单元格 (粗体矩形单元格, B13:D13); (3) 目标函数和约束左端项的代数定义 (虚线矩形单元格, D5:D9); (4) 提供说明的名称或符号的单元格. 规划求解只需要前 3 类信息. 第 4 类信息只是提高模型的可读性, 并没有其他的作用. 在电子表格中, 4 类信息的相对定位不必遵循图 2.12 所显示的布局. 例如, 定义目标函数和变量的单元格不必相邻, 也不必将它们放在问题下面. 重要的是我们知道它们在哪里, 以便可以被规划求解引用. 虽然如此, 最好使用类似于图 2.12 所建议的格式, 因为它使模型变得更易读.

规划求解如何链接电子表格中的数据呢? 首先我们用输入数据 (阴影的区域, B5:C9 和 F6:F9) 以及目标函数和变量 (粗体长形单元格, B13:D13) 构成, 提供与目标函数和约束左端项等价的“代数”定义, 然后我们将相关公式放在由虚线构成的矩形单元格 D5:D9 中适当的位置. 下表显示了原始线性规划函数以及所放置到的单元格位置:

	代数表达式	电子表格中的公式	输入单元格
目标 z	$5x_1+4x_2$	=B5*$B$13+C5*$C$13	D5
约束 1	$6x_1+4x_2$	=B6*$B$13+C6*$C$13	D6
约束 2	x_1+2x_2	=B7*$B$13+C7*$C$13	D7
约束 3	$-x_1+\ x_2$	=B8*$B$13+C8*$C$13	D8
约束 4	$0x_1+\ x_2$	=B9*$B$13+C9*$C$13	D9

① 其他著名的商业软件包包括 AIMMS、GAMS、LINGO、MPL、OPL Studio 和 Xpress-Mosel.



图 2.12　用 Excel 规划求解定义 Reddy Mikks 模型 (文件 solverRM1.xls)

实际上, 你只需在单元格 D5 处输入公式, 然后把它复制到其他的单元格 D6:D9 中. 为保证正确性, 必须用固定引用 \$B\$13 和 \$C\$13 表示 x_1 和 x_2. 对于较大的线性规划问题, 更有效的方法是在单元格 D5 中输入

$$= \text{SUMPRODUCT(B5:C5, \$B\$13:\$C\$13)}$$

然后把它复制到单元格 D6:D9 中.

线性规划模型的所有元素现在已准备好与“规划求解”链接. 从 Excel 的“工

具”菜单中, 选择“规划求解 (V)$\cdots$”[1] 打开如图 2.12 中间显示的“规划求解参数”对话框. 首先, 定义目标函数 z, 并由输入下列数据确定优化意义:

设置目标单元格 (E): $D$5

等于: ⊙ 最大值 (M)

可变单元格 (B): $B$13:$C$13

这些信息告诉规划求解, 由单元格 $B$13 到 $C$13 定义的变量由单元格 $D$5 中极大化目标函数确定.

下一步, 单击“**规划求解参数**”对话框中“ 添加 (A) ”键, 设置问题的约束. 计算机将显示“**添加约束**”对话框 (见图 2.12 底部), 你能够很容易地输入约束的元素 (左端项、不等式类型和右端项), 其形式如下[2]:

$D$6:$D$9 <= $F$6 : $F$9

在单元格范围内用键盘输入的另一种方便的方法是, 使单元格 D6:D9 加亮来输入约束的左端项, 接着再使单元格 F6:F9 加亮来输入约束的右端项. 相同的过程可以用在目标单元格.

剩下的约束只有非负限制, 为将非负限制加到模型中, 在“**添加约束**”对话框中单击“ 添加 ”, 输入

$B$13:$C$13 >= 0

输入非负约束的另一种方法是在“**规划求解参数**”对话框中, 单击“ 选项 (O) ”, 进入“**规划求解选项**”对话框 (见图 2.13), 然后在“ 假定非负 (G) ”前打勾 ☑. 当你还在“**规划求解选项**”对话框的状态下, 还需要在“ 采用线性模型 (M) ”前打勾 ☑.

一般而言, “**规划求解选项**”中的其余默认设置不必改动. 然而, 默认精度 0.000 001 对于某些问题而言可能被设置得太“高”了, 使得规划求解可能返回的消息是“规划求解找不到可行解”, 而事实上, 问题确实有可行解. 在这种情况下, 需要将精度调低到合适的精度. 如果还显示同样的信息, 则问题有可能是不可行的.

为了增强可读性, 可以用描述性的 Excel 区域的名称来代替单元格名称. 通过对所需要的单元格进行加亮, 在电子表格左边对话框的顶部键入区域名称, 然后按回车键来创建区域. 图 2.14(文件 solverRM2.xls) 概括了在此模型中使用的区域名称. 应该将文件 solverRM2.xls 与文件 solverRM1.xls 进行对比, 看一看区域是如何用在公式中的.

① 如果“规划求解”不出现在“工具”菜单下面, 单击该菜单下的“加载宏”并选择“规划求解”, 然后再单击“确定”.

② 你将注意到在“**添加约束**”对话框 (见图 2.12) 中, 中间的对话框指定了不等式的类型 (<= 和 >=), 它有另外的两种选项, int 和 bin[分别表示**整数**(integer) 和二进制数 (binary, 即 0 和 1)], 它们用在整数规划中, 用来限制变量取整数或取 0 和 1(见第 8 章).

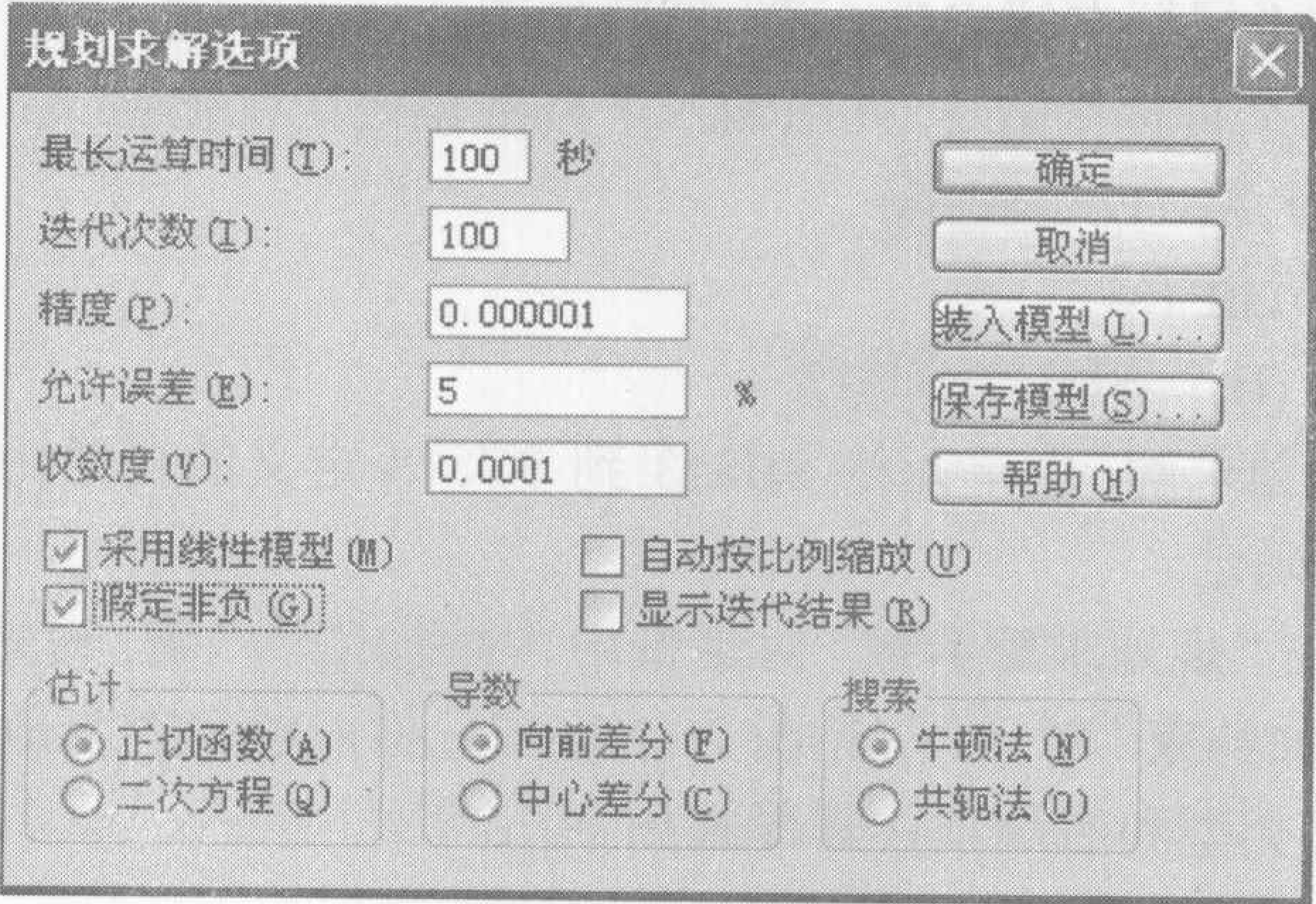


图 2.13　“规划求解选项”对话框

	A	B	C	D	E	F	G	H	I
1	Reddy Mikks Model								
2	Input data:							Range name	Cells
3		x1	x2					UnitsProduced	B13:C13
4		Exterior	Interior	Totals		Limits		UnitProfit	B5:C5
5	Objective	5	4	21				Totals	D6:D9
6	Raw material 1	6	4	24	<=	24		Limits	F6:F9
7	Raw material 2	1	2	6	<=	6		TotalProfit	D5
8	Market limit	-1	1	-1.5	<=	1			
9	Demand limit	0	1	1.5	<=	2			
10		>=0	>=0						
11	Output results:								
12		x1	x2	z					
13	Solution	3	1.5	21					

图 2.14　在 Excel 规划求解中区域名称的使用

为了求解问题, 单击“**规划求解参数**”对话框 (图 2.14) 中的“求解 (S)”. 这时出现一个新的对话框,“**规划求解结果**”, 它将会提供结果的状态. 如果模型设置是正确的, z 的最优值将出现在单元格 D5, 并且 x_1 和 x_2 的值将分别出现在单元格 B13 和 C13 中. 为了方便起见, 我们用单元格 D13 来显示 z 的最优值, 其方法

是在单元格 D13 中输入公式 =D5, 这样就可以在邻近的单元格中显示出完整的最优结果.

如果问题没有可行解, 规划求解将会显示明确的信息“规划求解找不到有用的解”. 如果最优目标值是无界的, 规划求解将会显示不太明确的信息“设置目标单元格的值未收敛”. 在任一种情况下, 这些信息都表明模型的公式中有错误, 这一点将在 3.5 节中予以讨论.

规划求解结果对话框将使你有机会要求 Excel 规划求解提供关于结果的更详细节, 包括重要的灵敏度分析报告. 我们将在 3.6.4 节讨论这些进一步的结果.

由规划求解得到的 Reddy Mikks 的解很简单. 而其他模型在它们能用方便的格式定义之前也许还需要“一点技巧性”. 套用此类模型的一类线性规划模型可以处理网络优化, 这将在第 6 章作讨论.

习题 2.4A

1. 修改图 2.12 所示的 Reddy Mikks 规划求解模型, 另增加一种具有防水功能的涂料. 每吨防水涂料需要的原料 1 和原料 2 分别是 0.5 吨和 0.75 吨. 新涂料的日需求在 0.5 吨至 1.5 吨之间. 其利润每吨是 3.5(单位: 1 000 美元).
2. 为下列问题构建 Excel 规划求解模型:
   (a) 例 2.2-2 的营养配方模型. (b) 习题 2.2A 中的第 16 题.
   (c) 例 2.3-1 城市更新模型.
   *(d) 例 2.3-2 套汇模型. (提示: 你将发现, 使用完整的货币转换矩阵要比仅用上三角元素要方便. 当然, 可以借助于适当的 Excel 公式产生下三角元素.)
   (e) 例 2.3-5 多周期生产–库存模型.

2.4.2 用 AMPL 解线性规划问题

本节[①] 对 AMPL 作一个简短的介绍. 在附录 A 的材料中, 提供了详细的 AMPL 语法内容, 本节的介绍, 以及贯穿在全书中的其他的 AMPL 介绍都会对它作适当的引用.

这里介绍 4 个例子: 前两个处理 AMPL 的基本概念, 后两个示范更高级的用法, 使 AMPL 的优点更明确.

Reddy Mikks 问题 —— 一个初等模型 AMPL 为线性规划建模提供了初等的句法格式. 图 2.15 为 Reddy Mikks 模型提供了清晰易懂的程序 (文件 amplRM1.txt). 所有的保留字都用黑体. 所有的其他名称都是由用户生成的. 目标函数和每一个约束必须给定明确的、用户生成的名称, 名称后紧跟着冒号 (:). 每条语句以分号 (;) 结束.

① 为方便起见, 在下载资料中, AMPL Optimization LLC 提供了 AMPL 学生版及相应的指导说明书. 进一步的更新资料可从网站 www.ampl.com 下载. AMPL 使用行命令, 并在 DOS(而不是 Windows) 环境下操作. 新近的 Windows 界面的 beta 版软件可以在 www.OptiRisk-Systems.com 找到.

```
var x1 >=0;
var x2 >=0;
maximize z: 5*x1+4*x2;
subject to
    c1: 6*x1+4*x2<=24;
    c2: x1+2*x2<=6;
    c3: -x1+x2<=1;
    c4: x2<=2;
solve;
display z,x1,x2;
```

图 2.15　Reddy Mikks 问题的初等 AMPL 模型 (文件 amplRM1.txt)

初等的 AMPL 模型太具体了, 具体到每改变一次数据, 都需要写出新的程序. 对于有成百 (甚至上千个) 变量和约束的实际问题, 这种写法太麻烦了. AMPL 为减轻这种困难, 将问题分成两个部分: (1) 一般模型, 用于表述任何数量的变量与约束的代数问题; (2) 特定数据, 用于驱动代数模型. 我们将用 Reddy Mikks 模型来演示 AMPL 的基本思想.

Reddy Mikks 问题 —— 一个代数模型　图 2.16 列出了模型 (文件 amplRM2.txt) 的语句. 文件必须是纯文本的 (ASCII). 注释语句前一定有 #, 并且可以出现在模型中的任何位置. 代码是区分大小写的, 并且所有的关键字 (极少数例外) 必须是小写字母. (详见附录 A.2.)

AMPL 中的代数模型将普通的 Reddy Mikks 问题看成下列一般格式:

$$\textbf{maximize } z: \sum_{j=1}^{n} c_j x_j$$

$$\textbf{subject to } restr_i: \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \quad i=1,2,\cdots,m$$

$$x_j \geqslant 0, \quad j=1,2,\cdots,n$$

它假定问题有 n 个变量和 m 个约束. 它给出了目标函数和 (任意) 约束 i, 其名称分别为 z 和 $restr_i$. 其余的参数 c_j, b_i 和 a_{ij} 是自明的.

模型以 `param` 语句开始, 声明 m, n, c, b 和 a_{ij} 为参数 (或常数), 其具体值是由模型的输入数据部分给出的. 它把 $c_j(j=1,2,\cdots,n)$ 翻译为 `c{1..n}`, 把 $b_i(i=1,2,\cdots,m)$ 翻译为 `b{1..m}`, 把 $a_{ij}(i=1,2,\cdots,m; j=1,2,\cdots,n)$ 翻译为 `a{1..m,1..n}`. 接下来, 变量 $x_j(j=1,2,\cdots,n)$ 结合非负限制, 由 `var` 语句定义为

```
var x{1..n}>=0;
```

如果将 `>=0` 从 x_j 的定义中除去, 则假定变量为无限制变量. `{}` 中的符号表示下标集, 其值在此之前已由 `parm` 或 `var` 定义.

在定义了所有的参数与变量后, 可以建立模型本身. 目标函数和每一个约束必

```
#---------------------------------------------algebraic model
param m;
param n;
param c{1..n};
param b{1..m};
param a{1..m,1..n};

var x{1..n}>=0;

maximize z: sum{j in 1..n}c[j]*x[j];
subject to restr{i in 1..m}:
         sum{j in 1..n}a[i,j]*x[j]<=b[i];
#---------------------------------------------specify model data
data;
param n:=2;
param m:=4;
param c:=1 5 2 4;
param b:=1 24  2 6  3 1  4 2;
param a:    1    2 :=
        1 6    4
        2 1    2
        3 -1   1
        4 0    1;
#---------------------------------------------solve the problem
solve;
display z, x;
```

图 2.16 带有输入数据的 Reddy Mikks 问题的 AMPL 模型 (文件 amplRM2.txt)

须带有明确的用户定义的名称, 名称紧跟冒号 (:). 在 Reddy Mikks 模型中, 目标函数由名称 `z:` 给出, 在它的前面是 `maximize`, 正像下面的 AMPL 语句所陈述的一样:

`maximize z: sum{j in 1..n} c[j]*x[j];`

语句是 maximize $z = \sum_{j=1}^{n} c_j x_j$ 的直接翻译 (用 : 替代 =). 注意, 方括号[ ]表示下标.

约束 i 是由根(root) 名称 `restr` 给出的, 并由集 `{1..m}` 做索引:

`restr{i in 1..m}:sum{j in  1..n}a[i,j]*x[j]<=b[i];`

语句是 $\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i$ 的直接翻译. 关键字 `subject to` 是可选项. 这个一般模型可以用于求解带有任何输入数据集 (表示任意 `m` 个约束和任意 `n` 个变量) 的任何问题.

data; 部分允许为特定的 Reddy Mikks 问题量身定做模型. 因此, param n:=2; 和 param m:=4; 告诉 AMPL, 问题有 2 个变量和 4 个约束. 注意, 必须使用复合运算符号 :=, 并且语句必须以关键字 param 开始. 对于单下标参数 c, 每一元素由下标 j 和由空格隔开后 c_j 的值来表示. 因此, 两个值 $c_1 = 5$ 和 $c_2 = 4$ 翻译成

```
param c:= 1 5 2 4;
```

参数 b 的数据可用类似的方法输入.

对于双下标参数 a, 顶行定义下标 j, 每行开始输入的是下标 i, 如

```
param a:   1    2 :=
      1    6    4
      2    1    2
      3   -1    1
      4    0    1;
```

在效果上, 数据 a_{ij} 作为二维矩阵来读, 其中它的行指定为 i, 列指定为 j. 注意, 分号只有在所有 a_{ij} 数据结束时需要.

模型和它的数据现已准备好. 命令 solve; 计算问题的解, 命令 display z, x; 提供计算结果.

为了执行模型, 首先要调用 AMPL(单击 AMPL 目录下的 ampl.exe). 在 ampl 的提示下, 输入如下model命令, 然后按回车键:

```
ampl: model AmplRM2.txt;
```

系统的输出则以如下形式呈现在屏幕上:

```
MINOS 5.5: optimal solution found.
2 iterations, objective = 21

z = 21
x[*]:=

1 = 3
2 = 1.5
```

最下面的 4 行是执行display z,x;的结果.

实际上, AMPL 允许将代数模型和数据分成两个独立的文件. 这种安排是可取的, 因为一旦模型建立完以后, 只需改动数据文件. (详见附录 A.2 尾部.) 本书并不选择分开模型与数据文件, 主要是为了节省篇幅.

套汇问题 简单的 Reddy Mikks 模型介绍了一些 AMPL 的基本原理. 例 2.3-2 更复杂的套汇模型为展示 AMPL 的更多能力提供了机会, 包括: (1) 附加在集元素上的条件; (2) if then else 语句的使用, 用来表示关系式; (3) 计算参数的使用; (4) 简单 print 语句的使用, 用来得到输出. 这些内容在附录 A 中也作了详细的讨

论.

图 2.17(文件 amplEx2.3-2.txt) 给出了套汇问题的 AMPL 代码. 模型具有普遍性, 可以用它极大化最终持有的名为 outCurrency 的任何货币 y (这些货币由初始量为 I、名为 nCurrency 的另一种货币转化而来). 还有, 任何货币的种类数 n, 可以包含在套汇过程中.

```
param inCurrency;                              # initial amount I
param outCurrency;                             # maximized holding y
param n;                                       #nbr of currencies
param r{i in 1..n,j in 1..n:i<=j};#above-diagonal rates
param I;                                       #initial amt of inCurrency
param maxTransaction{1..n};                    #limit on transaction amt

var x{i in 1..n,j in 1..n}>=0;                 #amt of i converted to j
var y>=0;                                      #max amt of outCurrency

maximize z: y;
subject to
 r1{i in 1..n,j in 1..n}:x[i,j]<=maxTransaction[i];
 r2{i in 1..n}:(if i=inCurrency then I else 0)+
   sum{k in 1..n}(if k<i then r[k,i] else 1/r[i,k])*x[k,i]=
     (if i=outCurrency then y else 0)+sum{j in 1..n}x[i,j];
#----------------------------------------input data
data;
param inCurrency=1;
param outCurrency=1;
param n:=5;
#            $    euro    pound     yen     KD
param r:     1    2       3         4       5:=
        1    1    .769    .625      105     .342    #$
        2    .    1       .813      137     .445    #euro
        3    .    .       1         169     .543    #pound
        4    .    .       .         1       .0032   #yen
        5    .    .       .         .       1;      #KD
param I:= 5;
param maxTransaction:=1 5 2 3 3 3.5 4 100 5 2.8;
#----------------------------------------Solution command
solve;
display z,y,x>file2.out;
print "rate of return =",trunc(100*(z-I)/I,4),"%">file2.out;
```

图 2.17 套汇问题 AMPL 模型 (文件 amplEx2.3-2.txt)

兑换率定义为

```
param r{i in 1..n,j in 1..n:i<=j};
```

通过在集{i in 1..n,j in 1..n}上使用强制条件 i<=j(前面有冒号), 定义仅给出对角线及其上方的元素. 由这个定义, 用倒数计算对角线下方的兑换率, 正如马上将要说明的那样.

变量 x_{ij}, 表示货币 i 兑换成货币 j 的数量, 定义为

```
var x{i in 1..n,j in 1..n}>=0;
```

模型有 2 组约束. 第 1 组约束, 以 r1 为根名称. 它使用如下语句, 限制任何货币兑换的交易量:

```
r1{i in 1..n,j in 1..n}: x[i,j]<=maxTransaction[i];
```

第 2 组约束, 以 r2 为根名称, 是下面的等式的转述:

$$\text{货币 } i \text{ 的输入} = \text{从货币 } i \text{ 得到的输出}$$

它的语句如下:

```
r2{i in 1..n}:
  (if i=inCurrency then I else 0)+
   sum{k in 1..n}(if k<i then r[k,i] else 1/r[i,k])*x[k,i]
       =(if i=outCurrency then y else 0)+sum{j in 1..n}x[i,j];
```

这种类型的约束最适合于使用特殊结构 if then else 来指定条件值. 在约束的左端, 表达式

```
(if i=inCurrency then I else 0)
```

说明在约束中, 对于输入货币 (i=inCurrency), 有外部的输入 I, 否则外部输入为零. 接下来, 表达式

```
sum{k in 1..n}(if k<i then r[k,i] else 1/r[i,k])*x[k,i]
```

计算从其他货币兑换成输入货币的输入资金. 如果回顾例 2.3-2, 你将注意到, 当 k<i 时, 兑换使用对角线上方元素的兑换率 r. 否则, 这行的倒数用在对角线下方的元素 (对角线元素为 1). 这正是 if then else 所做的. (详见附录 A.3.)

在约束 r2 的右端项中, if 表达式可以由类似的方法来解释, 即

```
(if i=outCurrency then y else 0)
```

说明对于 outCurrency, 外部输出是 y, 对其他所有情况为零.

我们可以提高约束 r2 的可读性, 方法是通过定义如下**计算参数**(computed parameter)(见附录 A.3) 来定义完整的兑换率表:

```
Param rate{k in 1..n,i in 1..n}
        =(if k<i then r[k,i] else 1/r[i,k])
```

在这种情况下, 约束 r2 变成

```
r2{i in 1..n}:
  (if i=inCurrency then I else 0)+ sum{k in 1..n}rate[k,i]*x[k,i]
      =(if i=outCurrency then y else 0)+sum{j in 1..n}x[i,j];
```

在 data; 部分, inCurrency 和 outCurrency 的每一个值等于 1, 这意味着

问题正在寻求初始数量为 500 万美元的情况下最大的美元输出. 在一般情况下, `inCurrency` 和 `outCurrency` 可以指定任何不同的货币. 例如, 设置 `inCurrency` 等于 2, `outCurrency` 等于 4, 表示在给出 500 万欧元的初始投资下, 极大化日元输出.

在 AMPL 中, `param r` 的未指定项用点 (.) 作标记. 然后这些值被图 2.17 所示的倒数, 或者在前面显示的计算参数 `rate` 所代替. 使用点的另一种做法是不必为对角线下的元素计算出数据并输入它们.

`display` 语句将输出送到文件 file2.out 中, 而不是默认时的输出到屏幕. `print` 语句对收益率作计算和舍位, 然后再将输出送到文件 file2.out 中. `print` 语句还可以用 `printf` 确定格式, 正像在任意的高级程序语言中一样 (详见附录 A.5.2).

最好要注意到, AMPL 中的输入数据不一定是模型中硬编码的, 因为它们可以从外部文件、电子表格、数据库来输入 (详见附录 A.5). 这一点在套汇模型中是至关重要的, 不稳定的兑换率必须经常在 10 秒内得到改变. 由于允许 AMPL 模型从自动更新兑换率的数据库中收到它的数据, 模型可以得到实时的最优结果.

公交车调度问题 例 2.3-8 的公交车调度问题提供了一个有趣的用 AMPL 建模的情形. 当然, 可以一直使用双下标参数, 它类似于 2.4.2 节介绍的 Reddy Mikks 模型 (图 2.16) 中的参数 a, 但在这种情况下, 这种方法可能是麻烦的. 相反, 我们可以利用约束的特殊结构, 使用条件表达式来表示它们的内在关系.

约束 (1) 的左端项是 $x_1 + x_m$, 其中 m 是一天 24 小时中总的周期数 (在前面的例子中 $= 6$). 对于其余的约束, 左端项所取的形式为 $x_{i-1} + x_i$, $i = 2, 3, \cdots, m$. 用 `if then else`(如我们在套汇问题中所做的), 所有 m 个约束, 正像图 2.18(文件

```
param m;
param  min_nbr_buses{1..m};
var x_nbr_buses{1..m} >= 0;
minimize tot_nbr_buses: sum {i in 1..m} x_nbr_buses[i];
subject to constr_nbr{i in 1..m}:
    if i=1 then
       x_nbr_buses[i]+x_nbr_buses[m]
    else
       x_nbr_buses[i-1]+x_nbr_buses[i]  >= min_nbr_buses[i];

data;
param m:=6;
param min_nbr_buses:= 1 4   2 8   3 10   4 7   5 12   6 4;

solve;
display tot_nbr_buses, x_nbr_buses;
```

图 2.18 例 2.3-8 公交车调度问题的 AMPL 模型 (文件 amplEx2.3-8.txt)

amplEx2.3-8.txt) 所显示的那样, 能够被一条语句简洁地表示. 这种表示优于约束左端项作为明确参数的定义.

AMPL 提供强大的编程能力. 例如, 输入或输出数据可以安全地由外部文件、电子表格和数据库读取或输出, 并且模型能够在一个较宽变化的、允许试验不同情景的选项下交互式地执行. 附录 A 中给出了详细资料. 另外, 贯穿全书的 AMPL 模型与附录 A 中的材料可以互相参照, 能够帮助你理解这些选项.

习题 2.4B

1. 在 Reddy Mikks 模型中, 假定生产第三类具有防水功能的涂料. 每吨防水涂料需要的原料 M1 和原料 M2 分别是 0.5 吨和 0.75 吨. 新涂料的日需求在 0.5 吨至 1.5 吨之间. 其利润每吨是 3.5(单位：1 000 美元). 修改 Excel 规划求解模型 solverRM2.xls 和 AMPL 模型 amplRM2.txt 适应于新的情形, 并且确定最优解. 结合每种修正, 对增加的工作量进行比较.
2. 为下列问题建立 AMPL 模型:
   (a) 例 2.2-2 的营养配方问题, 并求最优解. (b) 习题 2.3B 中的第 4 题.
   *(c) 习题 2.3D 中的第 7 题. (d) 习题 2.3G 中的第 7 题.
   (e) 习题 2.3G 中的第 9 题. *(f) 习题 2.3G 中的第 10 题.

参 考 文 献

Fourer, R., D. Gay, and B. Kernighan, *AMPL, A Modeling Language for Mathematical Programming*, 2nd ed., Brooks/Cole-Thomson, Pacific Grove, CA, 2003.

Kornbluth, J., and G. Salkin, *The Management for Corporate Financial Assets: Applications of Mathematical Programming Models*, Academic Press, London, 1987.

Shepard, R., D. Hartley, P. Hasman, L. Thorpe, and M. Bathe, *Applied Operations Research*, Plenum Press, New York, 1988.

Stark, R., and R. Nicholes, *Mathematical Programming Foundations for Design: Civil Engineering Systems*, McGraw-Hill, New York, 1972.

Laidlaw, C., *Linear Programming for Urban Development Plan Evaluation*, Praegers, London, 1972.

Lewis, T., "Personal Operations Research: Practicing OR on Ourselves," *Interfaces*, Vol. 26, No. 5, pp. 34 - 41, 1996.

William, H., *Model Building in Mathematical Programming*, 4th ed., Wiley, New York, 1994.

第3章 单纯形方法和灵敏度分析

本章导读 本章详细论述求解一般线性规划问题的单纯形方法. 还解释如何使用以单纯形为基础的灵敏度分析, 提供关于最优解重要的经济解释, 包含对偶价格和简约费用.

单纯形方法的计算非常单调乏味、反复重复, 尤其令人厌烦. 做这些计算时, 你应该寻找其几何解释, 也就是说, 单纯形方法尝试从解空间的一个角点移动到另一个更好的角点, 直到发现最优点为止. 为了在这方面为你提供帮助, TORA 的交互式用户引导模块 (立即反馈式) 允许你决定计算应该如何着手进行, 同时消除你冗长乏味的计算负担. 照这样下来, 你能够理解其概念, 而不必沉浸于计算的细节. 你尽管放心, 一旦你掌握了单纯形法的工作原理 (重要的是你确实理解这个概念), 计算机将完成冗长乏味的工作, 你将不再需要手工求解线性规划了.

教学经验告诉我, 当学生能够容易地完成冗长乏味的单纯形法计算以后, 一些学生还是不能说出为什么要做这些计算, 或结果的含义是什么. 为了帮助学生克服这个潜在的困难, 本章试图从原始问题解的角度出发, 强调每步迭代的解释.

学完本章内容后, 你将能够阅读和解释由商业软件提供的输出报告. 3.6 节描述这些报告在 AMPL、Excel规划求解和 TORA 中是如何产生的.

本章包含 1 个实际应用问题的概述、11 个例子、1 个 AMPL 模型、1 个 Excel规划求解模型、1 个 TORA 模型、107 个节后问题和 3 个案例. 案例放在本书 (下册) 的附录 E 中. AMPL/Excel/Solver/TORA 程序在下载资料的文件夹 ch3Files 中.

实际应用——心脏瓣膜生产的最优化

采用生物辅助手段制造不同规格的生物心脏瓣膜是用猪的心脏进行人工培植得到的. 从供应方来看, 猪的心脏不能按指定的大小规格来“生产”. 此外, 只有猪心脏的生物成分被处理过后, 才能够确定人工瓣膜的确切规格. 其结果是一些规格可能进贷过多, 而另一些规格可能是进贷不足. 当年建立了一个针对本问题的线性规划模型以减少进贷过多的规格并增加进贷不足的规格. 该模型的实际应用结果是, 在 1981 年这项研究完成的当年, 就节省了 1 476 000 多美元. 这项研究的详细资料见本书第 24 章中的案例 2.

3.1　等式形式的线性规划模型

可以通过对问题约束施加下面两个要求来方便单纯形法的计算:

(1) 所有的约束都是等式 (变量的非负限制除外), 并且具有非负的右端项.

(2) 所有变量是非负的.

这里施加的这两项要求主要是为了使单纯形方法标准化和简单化. 我们知道, 现在的所有商业软件包 (和 TORA) 都直接允许不等式约束、非负的右端项和无限制变量. 在单纯形法求解问题之前, 模型的任何必要的预处理都是在软件内部完成的.

3.1.1　将不等式转化为带有非负右端项的等式约束

在 ($\leqslant$) 约束中, 右端项可以被看作是关于资源可利用性限制的描述, 在这种情况下, 左端项表示由模型的活动 (变量) 对这些有限资源的用量. 因此, ($\leqslant$) 约束的右端项与左端项之间的差构成*未用的*或*松弛的*资源量.

为了把 ($\leqslant$) 不等式约束转换成等式约束, 在约束的左端, 增加非负的**松弛变量**(slack variable). 例如, 在 Reddy Mikks 模型中 (例 2.1-1), 相应于原料 M1 的约束给出如下:

$$6x_1 + 4x_2 \leqslant 24$$

定义 s_1 作为 M1 的松弛的或未用的量, 约束可以转换成如下等式约束:

$$6x_1 + 4x_2 + s_1 = 24, \quad s_1 \geqslant 0$$

接下来, 一个 ($\geqslant$) 约束设置了线性规划模型活动的下限, 因此可将不等式约束左端超出最小限制的量表示成*剩余*. 按如下方式完成从 ($\geqslant$) 到 ($=$) 的转换, 从不等式的左端项减去非负的**剩余变量** (surplus variable). 例如, 在营养配方模型中 (例 2.2-2), 表示最小饲料需求的约束是

$$x_1 + x_2 \geqslant 800$$

定义 S_1 作为剩余变量, 约束可以被转换成如下等式约束:

$$x_1 + x_2 - S_1 = 800, \quad S_1 \geqslant 0$$

仅剩下的要求就是让所得到的等式约束的右端项是非负的. 这个条件总是可以被满足, 必要时, 在得到方程的两端乘上 -1. 例如, 约束

$$-x_1 + x_2 \leqslant -3$$

的等价方程为

$$-x_1+x_2+s_1=-3,\ \ s_1\geqslant 0$$

现在, 两端乘上 -1, 转化为非负的右端项 (这就是我们要求的), 即

$$x_1-x_2-s_1=3$$

习题 3.1A

*1. 在 Reddy Mikks 模型 (例 2.2-1) 中, 考虑可行解 $x_1=3$ 吨, $x_2=1$ 吨. 为原料 M1 和 M2 确定相应的松弛变量的值.

2. 在营养配方模型 (例 2.2-2) 中, 确定由 500 磅玉米和 600 磅大豆粉构成饲料的剩余量.

3. 考虑不等式

$$10x_1-3x_2\geqslant -5$$

证明: 先在不等式两边乘上 -1 然后再转换不等式为等式约束与先变为等式然后再在两端乘上 -1 是相同的.

*4. 两种不同的产品 P1 和 P2 可以由一种或两种不同的机器 M1 和 M2 生产. 任何一种产品在任何一种机器上单位加工时间是相同的. 机器 M1 的日生产能力是 200 件 (或者是产品 P1, 或者是产品 P2, 或是两者的混合), 机器 M2 的日生产能力是 250 件. 车间的主管打算平衡这两种机器的生产进度, 使得一种机器生产的单位总数量与另一种机器生产的单位总数量相差在 5 件之内. 每件产品 P1 的利润是 10 美元, 产品 P2 是 15 美元. 将问题建立成等式约束形式的线性规划问题.

5. 说明如何将下列目标函数表示成等式形式:

$$\min\ z=\max\{|x_1-x_2+3x_3|,|-x_1+3x_2-x_3|\}$$
$$x_1,x_2,x_3\geqslant 0$$

(提示: $|a|\leqslant b$ 等价于 $-b\leqslant a\leqslant b$.)

6. 证明: m 个等式约束

$$\sum_{j=1}^{n}a_{ij}x_j=b_i,\ \ i=1,2,\cdots,m$$

等价于下列 $(m+1)$ 个不等式约束:

$$\sum_{j=1}^{n}a_{ij}x_j\leqslant b_i,\ \ i=1,2,\cdots,m$$

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{m}a_{ij}\right)x_j\geqslant\sum_{i=1}^{m}b_i$$

3.1.2 处理无限制变量

例 2.3-6 介绍了一个多周期生产平滑模型, 其中每个周期开始时的劳动力根据每个周期的需求作上下调整. 特别地, 如果 $x_i(\geqslant 0)$ 是周期 i 的劳动力数量, 则 $x_{i+1}(\geqslant 0)$ 是周期 $i+1$ 的劳动力数量, 可以被表示成

$$x_{i+1}=x_i+y_{i+1}$$

变量 y_{i+1} 必须无符号限制, 它允许 x_{i+1} 相对于 x_i 增加或减少, 分别取决于是雇用还是解雇工人.

正像我们不久将看到的那样, 单纯形方法的计算要求所有变量是非负的. 使用如下替换方法, 我们总可以满足这个要求.

$$y_{i+1} = y_{i+1}^- - y_{i+1}^+, \quad \text{其中} \quad y_{i+1}^- \geqslant 0, \ y_{i+1}^+ \geqslant 0$$

为了说明这种替换的工作原理, 假定在周期 1 中, 劳动力是 $x_1 = 20$ 名工人, 在周期 2 中, 劳动力将增加 5 名, 达到 25 名工人. 依据变量 y_2^- 和变量 y_2^+, 这将等价于 $y_2^- = 5$ 和 $y_2^+ = 0$, 或者 $y_2 = 5 - 0 = 5$. 类似地, 如果在周期 2 中, 劳动力减少到 16 名, 则我们有 $y_2^- = 0$ 和 $y_2^+ = 4$, 或者 $y_2 = 0 - 4 = -4$. 替换还允许劳动力有不作改变的可能性, 这可以通过令两个变量均为 0 来实现.

你或许会有疑问, y_2^- 和 y_2^+ 能否同时取正值. 直观地, 正像我们在例 2.3-6 中解释的那样, 这是不可能发生的, 因为这意味着在相同的时间, 我们既雇用工人也解雇工人. 这种直觉还得到数学证明的支持, 这个证明表明, 在任意的单纯形方法的解中, 这两个值同时取正值是不可能的.

习题 3.1B

1. McBurger 快餐店出售四分之一磅汉堡和干酪汉堡包. 四分之一磅汉堡使用 1/4 磅肉, 而干酪汉堡包仅用 0.2 磅肉. 餐馆以 200 磅肉开始一天的工作, 可以订购更多的肉, 但包括运输成本在内, 每磅肉的额外成本是 25 美分. 在一天结束时, 任何剩余肉均捐赠给慈善机构. McBurger 的利润是一份四分之一磅汉堡是 20 美分, 一份干酪汉堡包是 15 美分. McBurger 不指望在一天内的三明治销售量会超过 900 份. McBurger 应该计划一天应制作每种类型的三明治多少份? 用 TORA、Excel 规划求解和 AMPL 来求解此问题.

2. 某机器加工中心制造两种产品. 产品 1 和产品 2 的单位生产时间分别是 10 分钟和 12 分钟. 总的正常机器时间是每天 2 500 分钟. 在任意一天, 工厂可以生产 150 至 200 个单位的产品 1, 但不能超过 45 个单位的产品 2. 可以采用加班来满足需求, 但每分钟额外的成本是 0.5 美元. 假定产品 1 和产品 2 的单位利润分别是 6 美元和 7.5 美元, 将问题建立成一个线性规划模型, 然后用 TORA、Excel 规划求解或 AMPL 求出每种产品的最优生产水平和中心所需的加班时间.

***3.** JoShop 公司制造 3 种产品, 其单位利润分别是 2 美元、5 美元和 3 美元. 公司为 3 种产品的生产已经将 80 小时的劳动时间和 65 小时的机器时间编入预算. 每单位产品 1、产品 2 和产品 3 的劳动需求分别是 2 小时、1 小时和 2 小时. 相应的机器时间需求每单位分别是 1 小时、1 小时和 2 小时. JoShop 将编入预算的劳动时间和机器时间作为目标, 如果需要的话, 可以超过, 但每小时劳动的附加成本是 15 美元, 每小时的机器时间的附加成本是 10 美元. 将问题建立成线性规划模型, 并用 TORA、Excel 规划求解或 AMPL 确定最优解.

4. 在一个有若干个无限制变量的线性规划中, 变换 $x_j = x_j^- - x_j^+ (x_j^-, x_j^+ \geqslant 0)$ 将使非负变量的个数加倍. 一种较好的做法是, 我们可以用 $(k+1)$ 个非负变量替换 k 个无限制变量,

其替换方法是 $x_j = x'_j - w$, 其中 $x'_j, w \geqslant 0$. 用 TORA、Excel 规划求解或 AMPL 来说明, 对于如下线性规划, 两种方法产生的结果相同:

$$\begin{aligned} \max\ z &= -2x_1 + 3x_2 - 2x_3 \\ \text{s.t.}\ 4x_1 - x_2 - 5x_3 &= 10 \\ 2x_1 + 3x_2 + 2x_3 &= 12 \\ x_1 \geqslant 0, x_2, x_3 &\text{无限制} \end{aligned}$$

3.2 从图形解到代数解的转换

由 2.2 节线性规划的图解法所表达的思想奠定了代数单纯形法发展的基础. 图 3.1 画出了两种方法的类比. 在图解方法中, 解空间由表示约束的半空间描述; 而在单纯形法中, 解空间由 m 个同时成立的线性方程和 n 个非负变量表示.

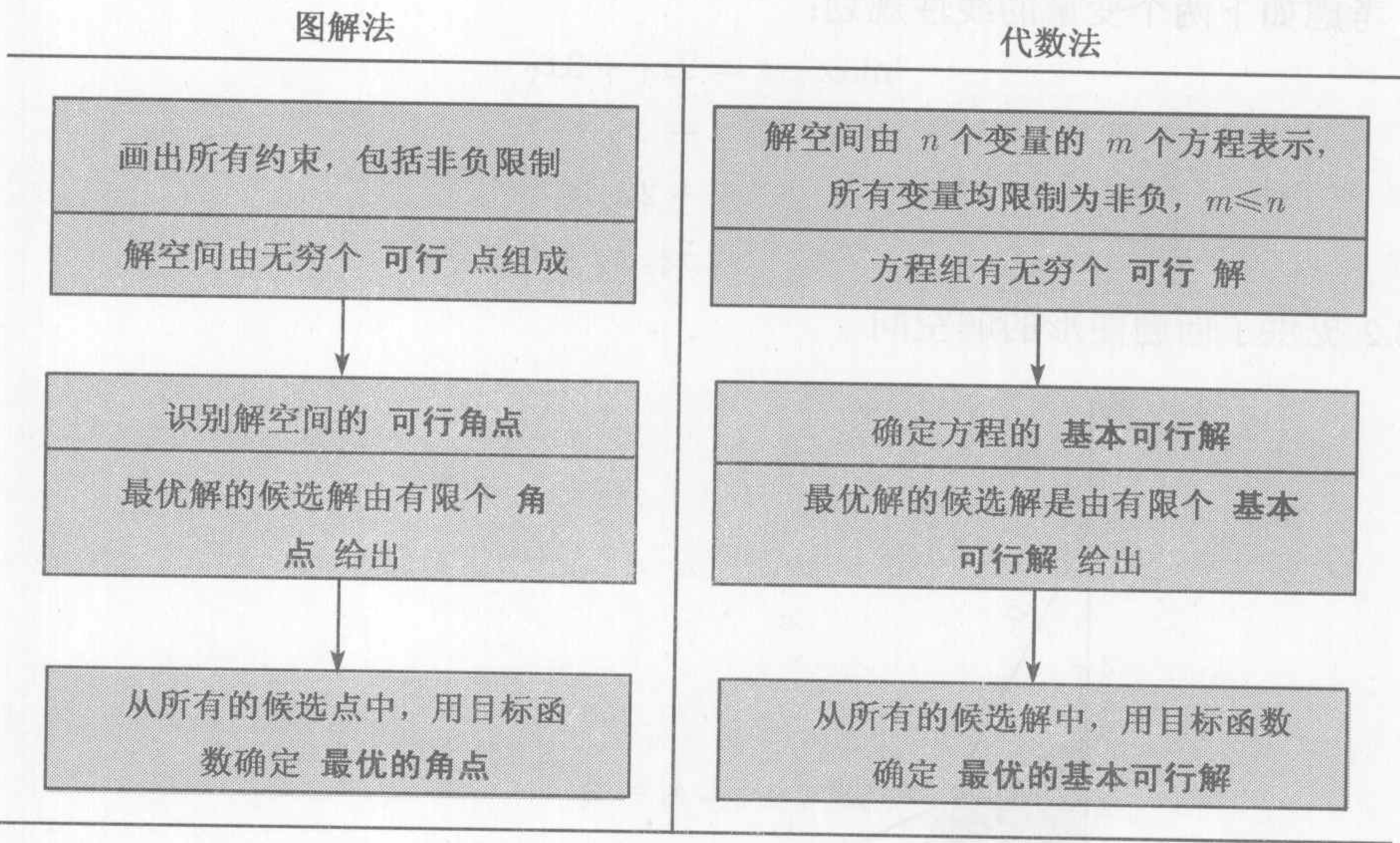


图 3.1 从图形解到代数解的转换

我们能够理解图形的解空间有无穷个解点的原因, 但如何能够从解空间的代数表示中得出类似的结论? 答案是, 在代数表示上, 方程的个数 m 总是小于等于变量的个数 n[①]. 如果 $m = n$, 并且方程是相容的, 则方程组只有唯一解; 但如果 $m < n$(它表示大多数线性规划), 并再次假定是相容的, 则方程组一定有无穷个解. 为了给出一个简单的解释, 例如方程 $x = 2$, 有 $m = n = 1$, 解显然是唯一的. 但是, 对于方程 $x + y = 1$ 而言, 有 $m = 1$ 和 $n = 2$, 它可以产生无穷个解 (在直线 $x + y = 1$ 上的任何点均是解).

① 如果方程的个数 m 大于变量的个数 n, 则至少有 $(m - n)$ 个方程必定是多余的.

在说明了如何以代数形式表示线性规划的解空间后, 用下列方法, 从联立的线性方程中确定最优值的候选点 (也就是角点):

角点的代数定义

在 $m \times n\ (m < n)$ 阶的方程组中, 如果令 $(n-m)$ 个变量等于 0, 然后求解其余的含 m 个变量的 m 个方程, 如果有唯一解, 则称相应的解为**基本解**(basic solution), 它一定对应解空间的一个 (可行或不可行) 角点. 这意味着角点的最大数目是

$$C_n^m = \frac{n!}{m!(n-m)!}$$

下面的例子演示上述过程.

例 3.2-1

考虑如下两个变量的线性规划:

$$\begin{aligned} \max \quad & z = 2x_1 + 3x_2 \\ \text{s.t.} \quad & 2x_1 + x_2 \leqslant 4 \\ & x_1 + 2x_2 \leqslant 5 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

图 3.2 提供了问题图形的解空间.

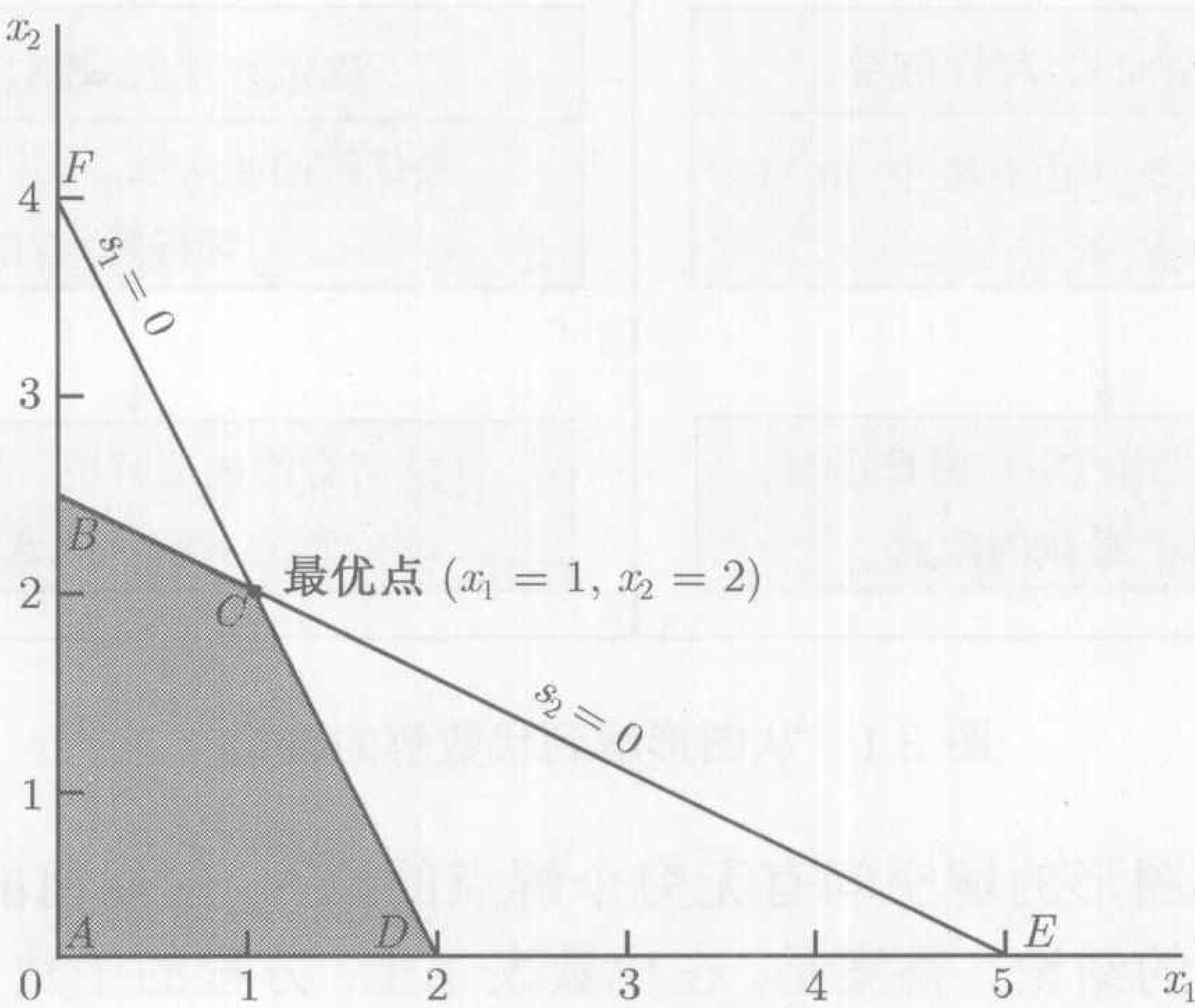


图 3.2　例 3.2-1 线性规划的解空间

从代数上看, 线性规划的解空间被表示为

$$\begin{aligned} 2x_1 + x_2 + s_1 &= 4 \\ x_1 + 2x_2 + s_2 &= 5 \\ x_1, x_2, s_1, s_2 &\geqslant 0 \end{aligned}$$

此方程组有 $m=2$ 个方程和 $n=4$ 个变量. 因此, 遵照给出的规则, 由代数式确定角点是令 $n-m=4-2=2$ 个变量为 0, 然后解其余的 $m=2$ 个变量. 例如, 如果令 $x_1=0$ 和 $x_2=0$, 方程提供了唯一的 (基本) 解:

$$s_1=4,\ s_2=5$$

这个解对应于图 3.2 中的点 A(自己检验 $s_1=4$ 和 $s_2=5$ 在点 A 处). 令 $s_1=0$ 和 $s_2=0$, 然后求解如下二元方程以确定另一个点:

$$\begin{aligned} 2x_1+\ \ x_2&=4 \\ x_1+2x_2&=5 \end{aligned}$$

它产生基本解 $(x_1=1, x_2=2)$, 它是图 3.2 中的点 C.

你或许会有疑问, 到底令哪些 $(n-m)$ 个变量为零才能对应一个特定的角点. 如果不利用图解法 (它仅对二维或三维变量有用) 的优点, 我们无法说哪些 $(n-m)$ 个零变量对应着哪个角点. 但这并不妨碍我们枚举解空间的全部角点. 只要我们考虑 $(n-m)$ 个变量为零时的全部组合, 并求解相应的方程. 一旦完成这项工作, 最优解就是达到最优目标值的基本可行解 (角点).

在前面的例子中, 我们有 $C_4^2=\frac{4!}{2!2!}=6$ 个角点. 看图 3.2, 我们可以立即认出 4 个角点 A, B, C, D. 那么, 剩下的两个点在哪儿? 事实上, 点 E 和 F 也是问题的角点, 但它们是不可行的, 因为它们不满足所有的约束. 这些不可行的角点不能成为最优解的候选点.

为概括从图形解到代数解的转换, 我们称 $(n-m)$ 个零变量为**非基变量** (nonbasic variable); 称余下的 m 个变量为**基变量** (basic variable), 它的解 (由解 m 个方程得到) 称为**基本解**. 下表提供了本例全部的基本解和非基本解.

非基 (零) 变量	基变量	基本解	相应的角点	可行否?	目标值 z
(x_1, x_2)	(s_1, s_2)	$(4, 5)$	A	是	0
(x_1, s_1)	(x_2, s_2)	$(4, -3)$	F	否	–
(x_1, s_2)	(x_2, s_1)	$(2.5, 1.5)$	B	是	7.5
(x_2, s_1)	(x_1, s_2)	$(2, 3)$	D	是	4
(x_2, s_2)	(x_1, s_1)	$(5, -6)$	E	否	–
$(\boldsymbol{s_1}, \boldsymbol{s_2})$	$(\boldsymbol{x_1}, \boldsymbol{x_2})$	$(\mathbf{1}, \mathbf{2})$	$\boldsymbol{C}$	**是**	**8(最优点)**

评注 我们从上面的计算能够看到, 当问题的大小增加后 (也就是 m 和 n 变大), 枚举所有角点的过程包含了巨量计算. 例如, 对于 $m=10$ 和 $n=20$, 必须求解 $C_{20}^{10}=184\ 756$ 个 10×10 阶的方程, 确实是一件令人吃惊的工作. 特别地, 当我们认识到在大多数实际应用情形中, (10×20) 的线性规划是一个小规模的问题, 因为

在应用中上百个甚至上千个变量和约束的问题并非不同寻常. 而单纯形方法通过借助于考查解空间中所有可能的基本可行解 (角点) 的一小部分, 极大地减轻了计算负担. 本质上, 单纯形方法利用一个智能的搜索过程, 用有效的方法查找最优角点的位置.

习题 3.2A

1. 考虑下面的线性规划:

$$\begin{aligned} \max\quad & z = 2x_1 + 3x_2 \\ \text{s.t.}\quad & x_1 + 3x_2 \leqslant 6 \\ & 3x_1 + 2x_2 \leqslant 6 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

(a) 将问题表示成等式约束形式.

(b) 确定问题所有的基本解, 并确定它们的可行性.

*(c) 在目标函数中使用直接替换来确定最优的基本可行解.

(d) 用图形验证, 在 (c) 中得到的解是线性规划的最优解. 因此, 得到如下结论: 在仅考虑基本可行解的情况下, 最优解能够由代数的方法确定.

*(e) 说明: 在图形的解空间中, 不可行的基本解是如何表示的.

2. 对下列每个线性规划, 通过枚举所有的基本可行解来确定最优解.

(a)
$$\begin{aligned} \max\quad & z = 2x_1 - 4x_2 + 5x_3 - 6x_4 \\ \text{s.t.}\quad & x_1 + 4x_2 - 2x_3 + 8x_4 \leqslant 2 \\ & -x_1 + 2x_2 + 3x_3 + 4x_4 \leqslant 1 \\ & x_1,\ x_2,\ x_3,\ x_4 \geqslant 0 \end{aligned}$$

(b)
$$\begin{aligned} \min\quad & z = x_1 + 2x_2 - 3x_3 - 2x_4 \\ \text{s.t.}\quad & x_1 + 2x_2 - 3x_3 + x_4 = 4 \\ & x_1 + 2x_2 + x_3 + 2x_4 = 4 \\ & x_1,\ x_2,\ x_3,\ x_4 \geqslant 0 \end{aligned}$$

***3.** 用代数方法证明, 下面线性规划所有的基本解都是不可行的.

$$\begin{aligned} \max\quad & z = x_1 + x_2 \\ \text{s.t.}\quad & x_1 + 2x_2 \leqslant 6 \\ & 2x_1 + x_2 \leqslant 16 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

4. 考虑下面的线性规划:

$$\begin{aligned} \max\quad & z = 2x_1 + 3x_2 + 5x_3 \\ \text{s.t.}\quad & -6x_1 + 7x_2 - 9x_3 \geqslant 4 \\ & x_1 + x_2 + 4x_3 = 10 \\ & x_1, x_3 \geqslant 0 \\ & x_2 \text{无限制} \end{aligned}$$

使用替换 $x_2 = x_2^- - x_2^+$, 将约束转换成等式形式. 证明: 基本解中不能同时包含 x_2^- 和 x_2^+.

5. 考虑下面的线性规划:

$$\max\quad z = x_1 + 3x_2$$

$$\begin{aligned} \text{s.t.} \quad x_1 + x_2 &\leqslant 2 \\ -x_1 + x_2 &\leqslant 4 \\ x_1\text{无限制}& \\ x_2 &\geqslant 0 \end{aligned}$$

(a) 确定问题所有的基本可行解.

(b) 在目标函数中用直接替换确定最优基本解.

(c) 用图解法求解, 并验证由 (b) 得到的解是最优的.

3.3 单纯形方法

与枚举线性规划问题 (正像我们在 3.2 节所做的) 的全部基本解 (角点) 相比, 单纯形法仅考查这些解中"少量特定的"解. 3.3.1 节描述这种方法的迭代本质, 3.3.2 节给出单纯形算法的计算细节.

3.3.1 单纯形方法的迭代本质

图 3.3 给出了例 3.2-1 线性规划的解空间. 正常情况下, 单纯形法从原点 (点 A)$x_1 = x_2 = 0$ 开始. 在这个初始点, 目标函数值 z 是零, 因此合乎逻辑的问题是, 能否在当前零值的基础上, 通过增加非基变量 x_1 和 (或) x_2 值来改进 (增加)z 的值. 考查如下目标函数来回答这个问题:

$$\max \ z = 2x_1 + 3x_2$$

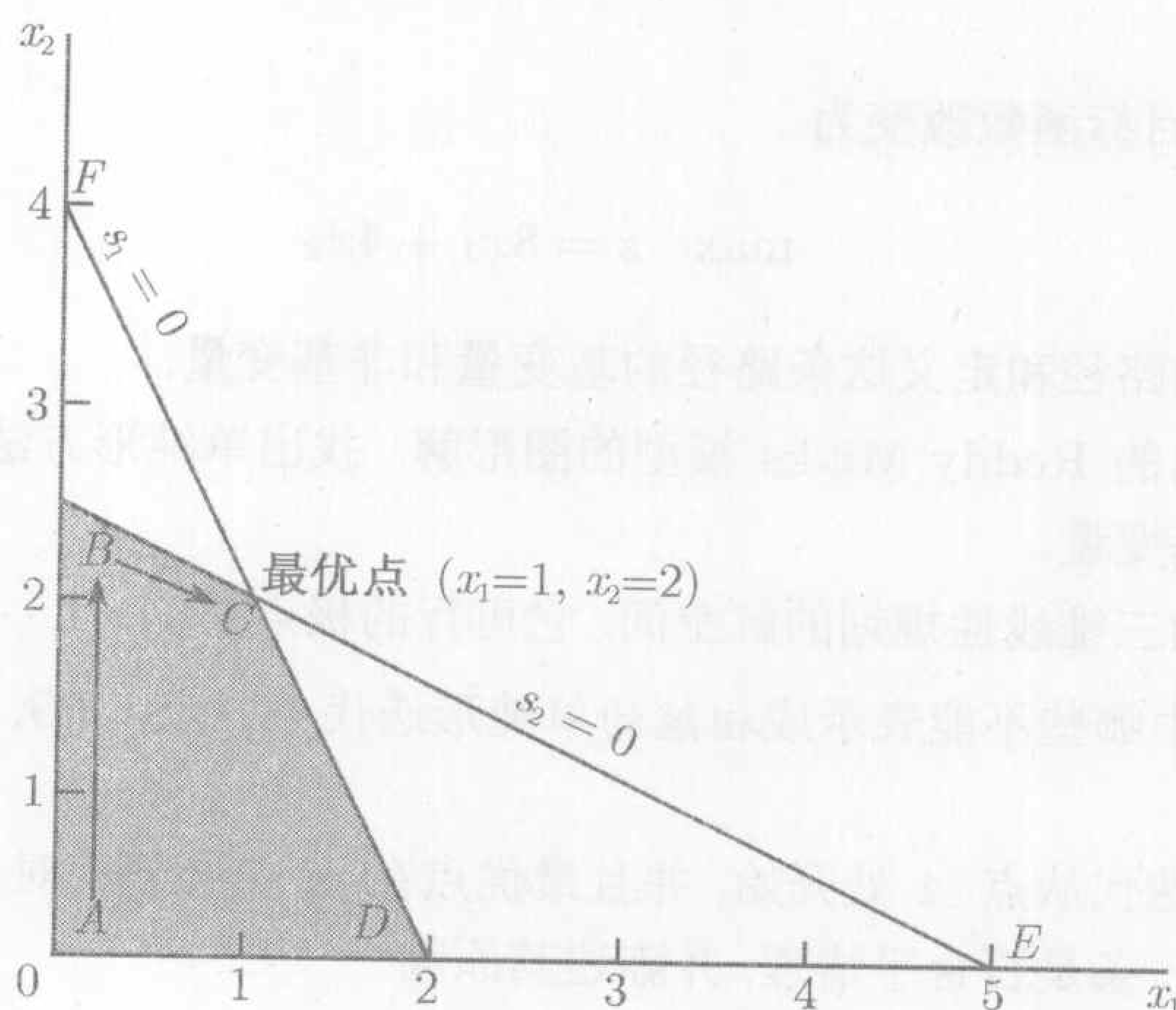


图 3.3 单纯形法的迭代过程

这个函数显示, 增加 x_1 或者增加 x_2 (或者同时增加) 将改进 z 的值. 单纯形方法的设计要求*每次增加一个变量*, 并且选择使 z 值有*最大改善率*的那个变量. 在前面的例子中, x_1 增加一个单位, 函数值增加 2; x_2 增加一个单位, 函数值增加 3. 这意味

着, z 值的改善率, 对于 x_1 是 2, 对于 x_2 是 3. 因此, 我们选择增加 x_2, 具有最大改善率的变量. 图 3.3 表明, 增加 x_2 的值直到角点 B(回顾一下, 在达不到角点 B 就停下来不是最优的, 因为最优点的候选点一定是角点). 在点 B, 单纯形方法这时再增加 x_1 的值, 到达改进的角点 C, 它是最优点. 因此, 单纯形算法的路径被定义为 $A \to B \to C$. 沿着路径的每个角点与一步**迭代** (iteration) 相对应. 要注意的是, 单纯形方法是沿着解空间的**边缘** (edge) 移动的, 这意味着该方法不能抄近路通过解空间, 从 A 直接走到 C.

借助于说明如何将点 A, B, C 表示成问题的基变量和非基变量, 我们需要使图形解转换成代数解. 下表概括了这些表示方法:

角点	基变量	非基 (零) 变量
A	s_1, s_2	x_1, x_2
B	s_1, x_2	x_1, s_2
C	x_1, x_2	s_1, s_2

注意, 在基变量和非基变量中的变化模式随着解沿路径 $A \to B \to C$ 的移动而改变. 从 $A \to B$, 在 A 处的非基变量 x_2 变成 B 处的基变量, 并且在 A 处的基变量 s_2 变成在 B 处的非基变量. 在单纯形方法的术语中, 我们称 x_2 为**进基变量**(entering variable, 因为它进入基本可行解), 称 s_2 为**离基变量**(leaving variable, 因为它离开基本可行解). 用类似的方法, 在点 B, x_1 进基, s_1 离基, 因此, 走到了点 C.

习题 3.3A

1. 在图 3.3 中, 假设目标函数改变为

$$\max \quad z = 8x_1 + 4x_2$$

确定单纯形方法的路径和定义这条路径的基变量和非基变量.

2. 考虑由图 2.2 给出的 Reddy Mikks 模型的图形解. 找出单纯形方法的路径和定义这条路径的基变量和非基变量.

***3.** 考虑图 3.4 给出的三维线性规划的解空间. 它可行的极点是 $A, B, \cdots, J$.

(a) 下列角点对中哪些不能表示成相继的单纯形迭代: $(A, B), (B, D), (E, H)$ 和 (A, I)? 解释其原因.

(b) 假定单纯形迭代从点 A 处开始, 并且最优点在 H 处出现. 对于单纯形算法, 指出下列路径的每一条是否合乎情理, 并陈述其原因.

(i) $A \to B \to G \to H$.　　**(ii)** $A \to E \to I \to H$.

(iii) $A \to C \to E \to B \to A \to D \to G \to H$.

4. 在图 3.4 中的解空间, 所有约束是 $\leqslant$ 类型, 并且所有变量 x_1, x_2, x_3 是非负的. 假定 s_1, s_2, s_3, $s_4 (\geqslant 0)$ 分别是相应约束的松弛变量, 其约束分别表示平面 $CEIJF$、$BEIHG$、$DFJHG$ 和 IJH. 找出与解空间每个可行极点相对应的基变量和非基变量.

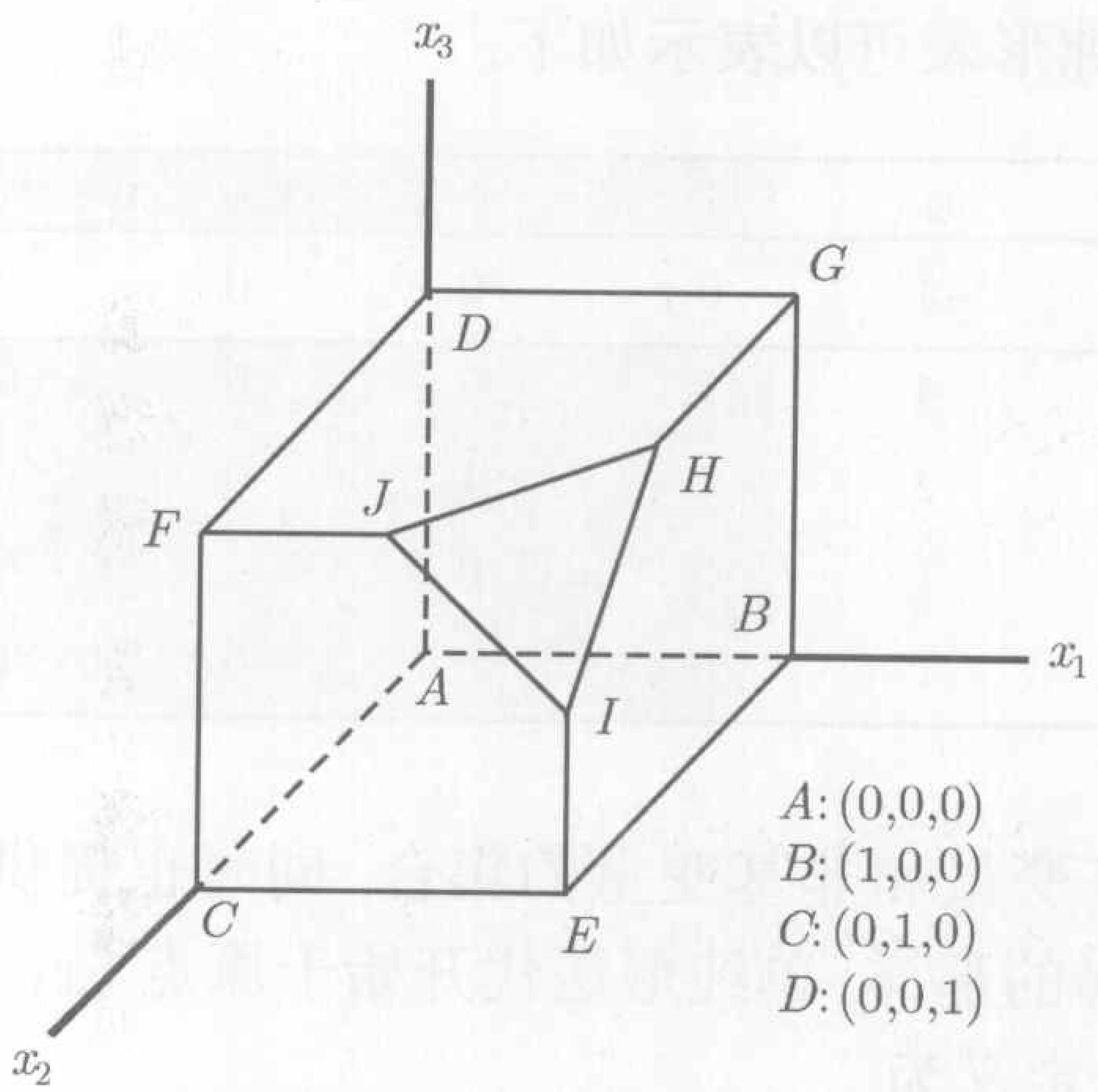


图 3.4 习题 3.3A 第 3 题的解空间

5. 考虑图 3.4 中的解空间, 这里单纯形算法从点 A 处开始. 在首次迭代中, 结合其值和 z 值的改进, 对下列目标函数, 确定进基变量.

*(a) $\max\ z = x_1 - 2x_2 + 3x_3$ (b) $\max\ z = 5x_1 + 2x_2 + 4x_3$

(c) $\max\ z = -2x_1 + 7x_2 + 2x_3$ (d) $\max\ z = x_1 + x_2 + x_3$

3.3.2 单纯形算法的计算细节

本节提供一步单纯形迭代的计算细节, 包括确定进基变量和离基变量的准则, 以及达到最优解时的终止计算准则, 通过一个数值例子给出解释.

例 3.3-1

我们用 Reddy Mikks 模型 (例 2.2-1) 解释单纯形方法的细节. 将问题表示成等式约束形式如下:

$$\max\ z = 5x_1 + 4x_2 + 0s_1 + 0s_2 + 0s_3 + 0s_4$$

$$\begin{aligned}
\text{s.t.}\quad 6x_1 + 4x_2 + s_1 \qquad\qquad &= 24 \quad (\text{原料M1})\\
x_1 + 2x_2 \qquad + s_2 \qquad &= 6 \quad (\text{原料M2})\\
-x_1 + x_2 \qquad\qquad + s_3 &= 1 \quad (\text{市场限制})\\
x_2 \qquad\qquad\qquad + s_4 &= 2 \quad (\text{需求限制})
\end{aligned}$$

$$x_1, x_2, s_1, s_2, s_3, s_4 \geqslant 0$$

变量 s_1, s_2, s_3, s_4 分别对应于约束的松弛变量.

下面, 写出目标方程如下:

$$z - 5x_1 - 4x_2 = 0$$

在这种方法中, 初始单纯形表可以表示如下:

基	z	x_1	x_2	s_1	s_2	s_3	s_4	解	
z	1	−5	−4	0	0	0	0	0	z 行
s_1	0	6	4	1	0	0	0	24	s_1 行
s_2	0	1	2	0	1	0	0	6	s_2 行
s_3	0	−1	1	0	0	1	0	1	s_3 行
s_4	0	0	1	0	0	0	1	2	s_4 行

表的设计指定了基变量和非基变量的集合, 同时也提供了相应于初始迭代的解. 正如在 3.3.1 节解释的那样, 单纯形迭代开始于原点 $(x_1, x_2) = (0, 0)$, 相应的非基变量和基变量的集合定义为

$$\text{非基 (零) 变量: } (x_1, x_2)$$

$$\text{基变量: } (s_1, s_2, s_3, s_4)$$

令非基变量 $(x_1, x_2) = (0, 0)$, 并注意到在表中 z 的系数和基变量 (s_1, s_2, s_3, s_4) 特定的 0-1 排列, 可以立刻得到 (不需要任何计算) 下面的解:

$$z = 0$$
$$s_1 = 24$$
$$s_2 = 6$$
$$s_3 = 1$$
$$s_4 = 2$$

该信息可由表中最左端**基**列所列出的基变量和最右端**解**列所列出的对应值来显示. 在效果上, 该表通过指定它的基变量及其值, 定义了当前的角点和目标函数 z 相应的值. 记住, 非基变量 (那些没有放在**基**列中的变量) 总是等于零.

初始解是最优的吗? 目标函数 $z = 5x_1 + 4x_2$ 表明可以通过增加 x_1 或 x_2 来改进这个解. 根据 3.3.1 节中的讨论, 选择具有*最正*(most positive) 系数的变量 x_1 作为*进基变量*. 这等价于, 因为单纯形表将目标函数表示为 $z - 5x_1 - 4x_2 = 0$, 则进基变量将对应于目标方程中那些*最负*(most negative)[①]系数的变量. 这个规则被称为**最优性条件** (optimality condition).

从单纯形表中确定离基变量的技巧是, 计算方程的右端项 (**解**列) 与相应的在进基变量 x_1 下方的约束系数的*非负比*, 如下表所示.

---

① 指绝对值最大的负数.—— 译者注

基	进基 x_1	解	比 (或截距)
s_1	6	24	$x_1 = \frac{24}{6} = 4$ ← 最小值
s_2	1	6	$x_1 = \frac{6}{1} = 6$
s_3	−1	1	$x_1 = \frac{1}{-1} = -1$ (不考虑)
s_4	0	2	$x_1 = \frac{2}{0} = \infty$ (不考虑)
结论: x_1 进基, s_1 离基			

最小非负比自动识别当前基变量 s_1 作为离基变量, 并指定进基变量 x_1 的新值为 4.

所计算出的比如何确定离基变量和进基变量的值? 图 3.5 显示, 由计算得到的这些比值实际上是这些约束关于进基变量 (x_1) 轴的截距. 可以看到, x_1 的值必须增加到 4, 达到角点 B, 它是关于 x_1 轴最小的非负截距. 再增加而超过 B 是不可行的. 在点 B, 当前与约束①相应的基变量 s_1 设定为零值, 变成*离基变量*. 相应于比值计算的规则被称为**可行性条件** (feasibility condition), 因为它保证新解的可行性.

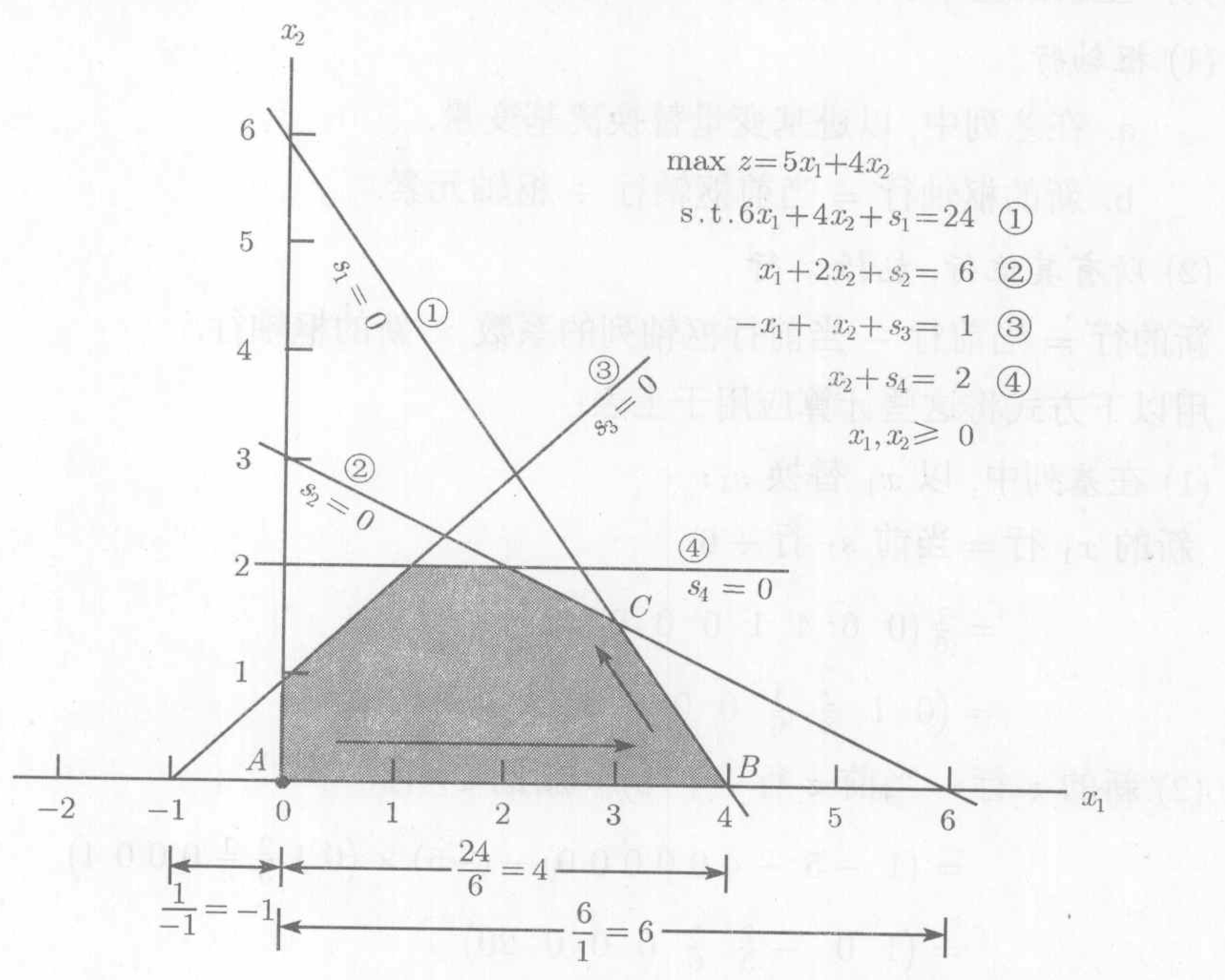


图 3.5 在 Reddy Mikks 模型中, 单纯形方法比的图形解释

新解点 B 由进基变量 x_1 和离基变量 s_1 在单纯形表中的"交换"来确定, 产生下列一组非基变量和基变量如下:

在点 B 处的非基 (零) 变量: (s_1, x_2)

在点 B 处的基变量: (x_1, s_2, s_3, s_4)

交换过程基于**高斯–若尔当行运算** (Gauss-Jordan row operation). 它认定进基变量所在的列为**枢轴列** (pivot column), 离基变量所在的行为**枢轴行** (pivot row). 枢轴列与枢轴行交叉位置的元素称为**枢轴元素** (pivot element). 下表重新陈述了初始单纯形表, 其中突出显示了枢轴行和枢轴列.

	基	z	x_1 (进基 ↓)	x_2	s_1	s_2	s_3	s_4	解	
	z	1	−5	−4	0	0	0	0	0	
离基 ←	s_1	0	**6**	4	1	0	0	0	24	枢轴行
	s_2	0	1	2	0	1	0	0	6	
	s_3	0	−1	1	0	0	1	0	1	
	s_4	0	0	1	0	0	0	1	2	
			枢轴列							

为产生新的基本解, 高斯–若尔当计算需要产生包括两种类型的新的基本解.

(1) 枢轴行

a. 在基列中, 以进基变量替换离基变量.

b. 新的枢轴行 = 当前枢轴行 ÷ 枢轴元素.

(2) 所有其他行, 包括 z 行

新的行 = 当前行 − 当前行枢轴列的系数 × 新的枢轴行.

用以下方式将这些计算应用于上表:

(1) 在基列中, 以 x_1 替换 s_1:

新的 x_1 行 = 当前 s_1 行 ÷ 6

$$= \tfrac{1}{6}(0\ \ 6\ \ 4\ \ 1\ \ 0\ \ 0\ \ 0\ \ 24)$$

$$= (0\ \ 1\ \ \tfrac{2}{3}\ \ \tfrac{1}{6}\ \ 0\ \ 0\ \ 0\ \ 4)$$

(2) 新的 z 行 = 当前 z 行 − (−5) × 新的 x_1 行

$$= (1\ \ -5\ \ -4\ 0\ 0\ 0\ 0\ 0\ 0) - (-5) \times (0\ 1\ \tfrac{2}{3}\ \tfrac{1}{6}\ 0\ 0\ 0\ 4)$$

$$= (1\ \ 0\ \ -\tfrac{2}{3}\ \ \tfrac{5}{6}\ \ 0\ \ 0\ \ 0\ \ 20)$$

(3) 新的 s_2 行 = 当前 s_2 行 − (1) × 新的 x_1 行

$$= (0\ 1\ 2\ 0\ 1\ 0\ 0\ 6) - (1) \times (0\ 1\ \tfrac{2}{3}\ \tfrac{1}{6}\ 0\ 0\ 0\ 4)$$

$$= (0\ \ 0\ \ \tfrac{4}{3}\ \ -\tfrac{1}{6}\ \ 1\ \ 0\ \ 0\ \ 2)$$

(4) 新的 s_3 行 = 当前 s_3 行 $-(-1)\times$ 新的 x_1 行

$$=(0\ -1\ 1\ 0\ 0\ 1\ 0\ 1)-(-1)\times(0\ 1\ \tfrac{2}{3}\ \tfrac{1}{6}\ 0\ 0\ 0\ 4)$$

$$=(0\ \ 0\ \ \tfrac{5}{3}\ \ \tfrac{1}{6}\ \ 0\ \ 1\ \ 0\ \ 5)$$

(5) 新的 s_4 行 = 当前 s_4 行 $-(0)\times$ 新的 x_1 行

$$=(0\ 0\ 1\ 0\ 0\ 0\ 1\ 2)-(0)\times(0\ 1\ \tfrac{2}{3}\ \tfrac{1}{6}\ 0\ 0\ 0\ 4)$$

$$=(0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\ \ 1\ \ 2)$$

新的基本解是 (x_1, s_2, s_3, s_4), 因此新的单纯形表变为:

	基	z	x_1	↓ x_2	s_1	s_2	s_3	s_4	解
	z	1	0	$-\frac{2}{3}$	$\frac{5}{6}$	0	0	0	20
	x_1	0	1	$\frac{2}{3}$	$\frac{1}{6}$	0	1	0	4
←	s_2	0	0	$\frac{4}{3}$	$-\frac{1}{6}$	1	0	0	2
	s_3	0	0	$\frac{5}{3}$	$\frac{1}{6}$	0	1	0	5
	s_4	0	0	1	0	0	0	1	2

观察到新表与初始表有相同的性质. 当我们置新的非基变量 x_2 和 s_1 为 0, 解列自动生成新的基本解 $(x_1=4, s_2=2, s_3=5, s_4=2)$. 这种表的"状况"是高斯–若尔当行运算应用的结果. 相应新的目标函数值是 $z=20$, 它与下面的公式相一致:

新的 z = 原来的 z + 新的 x_1 的值 $\times$ 它的目标系数

$$=0+4\times 5=20$$

在最后的表中, 最优性条件表明, x_2 是进基变量. 由可行性条件可得下表:

基	进基 x_2	解	比 值
x_1	$\frac{2}{3}$	4	$x_2=4\div\frac{2}{3}=6$
s_2	$\frac{4}{3}$	2	$x_2=2\div\frac{4}{3}=1.5$ (最小值)
s_3	$\frac{5}{3}$	5	$x_2=5\div\frac{5}{3}=3$
s_4	1	2	$x_2=2\div 1=2$

因此, s_2 离开基本解, 并且 x_2 的新值是 1.5. 相应增加的 z 值是 $\frac{2}{3}x_2=\frac{2}{3}\times 1.5=1$, 它产生新的 $z=20+1=21$.

在基列中, 用进基变量 x_2 替换 s_2, 应用高斯–若尔当行运算:

(1) 新的枢轴 x_2 行 = 当前 s_2 行 $\div\frac{4}{3}$;

(2) 新的 z 行 = 当前 z 行 $-(-\frac{2}{3})\times$ 新的 x_2 行;

(3) 新的 x_1 行 = 当前 x_1 行 $-(\frac{2}{3})\times$ 新的 x_2 行;

(4) 新的 s_3 行 = 当前 s_3 行 $-(\frac{5}{3})\times$ 新的 x_2 行;

(5) 新的 s_4 行 = 当前 s_4 行 $-(1)\times$ 新的 x_2 行.

这些计算产生下表:

基	z	x_1	x_2	s_1	s_2	s_3	s_4	解
z	1	0	0	$\frac{3}{4}$	$\frac{1}{2}$	0	0	21
x_1	0	1	0	$\frac{1}{4}$	$-\frac{1}{2}$	0	0	3
x_2	0	0	1	$-\frac{1}{8}$	$\frac{3}{4}$	0	0	$\frac{3}{2}$
s_3	0	0	0	$\frac{3}{8}$	$-\frac{5}{4}$	1	0	$\frac{5}{2}$
s_4	0	0	0	$\frac{1}{8}$	$-\frac{3}{4}$	0	1	$\frac{1}{2}$

基于最优性条件, z 行中相应于非基变量 s_1 和 s_2 的系数没有一个是负的. 因此, 最后一张表是最优的.

可以用下列方式从单纯形表中读出最优解. 基列中变量的最优值由右端项解列给出, 并且可以被解释成:

决策变量	最优值	建　议
x_1	3	日生产 3 吨外墙涂料
x_2	$\frac{3}{2}$	日生产 1.5 吨内墙涂料
z	21	日利润是 21 000 美元

可以验证, 将 $s_1 = s_2 = 0$, $s_3 = \frac{5}{2}$, $s_4 = \frac{1}{2}$ 代入方程替换出的 x_1 与 x_2 的值与 x_1 与 x_2 的已知值是一致的.

最优解还给出了资源的状况. 一种资源被认为是**匮乏的** (scarce), 如果模型的活动 (变量) 使用了全部资源; 否则资源是**充裕的**(abundant). 这个信息可以通过检验相应的表示资源约束的松弛变量的值, 由最优单纯形表得到确定. 如果松弛变量的值是零, 则资源全部用完, 因此, 它们是匮乏的. 否则, 一个正的松弛变量表明资源是充裕的. 下表将模型的约束作了分类:

资　源	松弛变量的值	状况
原料 M1	$s_1 = 0$	匮乏
原料 M2	$s_2 = 0$	匮乏
市场限制	$s_3 = \frac{5}{2}$	充裕
需求限制	$s_4 = \frac{1}{2}$	充裕

评注 单纯形表提供了有价值的附加信息, 包括:

(1) *灵敏度分析*, 确定使当前解保持不变的条件;

(2) *后最优分析*, 确定当模型的数据变化后, 如何找出新的最优解.

3.6 节介绍灵敏度分析. 后最优分析详见第 4 章的内容.

TORA 程序

高斯–若尔当计算单调乏味、繁琐冗长，所有这些是令人厌烦的. 但是，这些都还是次要的，因为在实际中这些计算可由计算机完成. 重要的是你要了解单纯形法如何工作. TORA 交互式用户引导(立即反馈式) 选项可以在这一点上给你帮助，因为在单纯形法中，它允许你决定计算过程，而没有执行高斯–若尔当计算的负担. 为了用 TORA 求解 Reddy Mikks 问题，输入模型，而后从 SOLVE/MODIFY 菜单中，选择 Solve ⇒ Algebraic ⇒ Iterations ⇒ All-Slack. (All-Slack 选择表明，初始基本解仅由松弛变量组成. 余下的选项将在 3.4 节、4.3 节和 13.4.2 节介绍.) 接下来，单击 Go To Output Screen. 单击 Next Iteration 或 All Iterations 之后，可以产生一次或所有的迭代. 如果选择每次产生一次迭代，那么单击相应的列与行的标题，你就能交互式地指定进基和离基变量. 如果你的选择是正确的，其列变绿，其行变红. 否则，将弹出一个出错信息.

3.3.3 单纯形法的总结

到现在为止，我们已经处理了最大化的情形. 在最小化的问题中，最优性条件与最大化情形的规则正好相反，要求选择在目标方程中目标系数是最正的非基变量为进基变量. 这是因为极大化 z 等价于极小化 $(-z)$. 至于选择离基变量的可行性条件，其规则保持不变.

最优性条件 在极大化 (极小化) 问题中，进基变量是 z 行中具有最负 (最正) 系数的非基变量. 如有多个可任选其一. 当非基变量的所有 z 行系数是非负的 (非正的) 时，迭代达到最优值.

可行性条件 对于极大化和极小化问题，离基变量都是具有最小非负比 (带有严格的正分母) 的基变量. 如有多个可任选其一.

高斯–若尔当行运算

(1) 枢轴行

a. 在基列中，用进基变量替换离基变量.

b. 新的枢轴行 = 当前枢轴行 ÷ 枢轴元素.

(2) 包括 z 的所有其他行

新行 = 当前行 − 枢轴列系数 × 新的枢轴行.

单纯形方法的步骤是:

第 1 步 确定初始基本可行解.

第 2 步 用最优性条件选择一个进基变量. 如果没有进基变量，停止计算；上一个解就是最优的. 否则，转到第 3 步.

第 3 步 用可行性条件选择离基变量.

第 4 步 用适当的高斯–若尔当行运算确定新的基本解. 转到第 2 步.

习题 3.3B

1. 设计这个问题是为了加强你对单纯形可行性条件的理解. 在例 3.3-1 的初始表中, 我们用最小 (非负) 比检验来确定离基变量. 这样一个条件保证了基变量新值均不会成为负值 (正如线性规定义所规定的). 为了证实这一点, 迫使 s_2(不是 s_1) 离开基本解. 现在, 看一下导出的单纯形表, 你将注意到, s_1 呈现出负值 $(=-12)$, 这意味着新的解是不可行的. 如果使用最小比可行性条件, 这种情况将永远不会发生.

2. 考虑下列约束集:

$$\begin{aligned}
x_1+2x_2+2x_3+4x_4 &\leqslant 40\\
2x_1-x_2+x_3+2x_4 &\leqslant 8\\
4x_1-2x_2+x_3-x_4 &\leqslant 10\\
x_1,x_2,x_3,x_4 &\geqslant 0
\end{aligned}$$

对于下列目标函数求解问题.

(a) $\max\ z=2x_1+x_2-3x_3+5x_4$ **(b)** $\max\ z=8x_1+6x_2+3x_3-2x_4$

(c) $\max\ z=3x_1-x_2+3x_3+4x_4$ **(d)** $\min\ z=5x_1-4x_2+6x_3-8x_4$

***3.** 考虑下列方程组:

$$\begin{aligned}
x_1+2x_2-3x_3+5x_4+x_5 \qquad\qquad\qquad &=4\\
5x_1-2x_2 \qquad +6x_4 \qquad +x_6 \qquad\qquad &=8\\
2x_1+3x_2-2x_3+3x_4 \qquad\qquad +x_7 \qquad &=3\\
-x_1 \qquad +x_3-2x_4 \qquad\qquad\qquad +x_8 &=0\\
x_1,x_2,\cdots,x_8 &\geqslant 0
\end{aligned}$$

令 x_5, x_6, x_7, x_8 是已知的基本可行解. 假定 x_1 变成基变量. 在已知基变量中哪一个变成非基变量而取零值, 以保证其余所有变量保持非负? 在新的解中 x_1 的值是多少? 对于 x_2, x_3, x_4 重复这个过程.

4. 考虑下面的线性规划:

$$\begin{aligned}
\max\quad & z=x_1\\
\text{s.t.}\quad & 5x_1+x_2 =4\\
& 6x_1+x_3=8\\
& 3x_1+x_4=3\\
& x_1,x_2,x_3,x_4\geqslant 0
\end{aligned}$$

(a) 用观察的方法 (不用高斯–若尔当行运算) 求解问题, 从单纯形方法基本解的角度来证明答案是正确的.

(b) 重复 (a), 假定目标函数是求 $\min\ z=x_1$.

5. 用观察的方法求解下面的问题, 并从单纯形方法基本解的角度证明其求解方法是正确的.

$$\begin{aligned}
\max\quad & z=5x_1-6x_2+3x_3-5x_4+12x_5\\
\text{s.t.}\quad & x_1+3x_2+5x_3+6x_4+3x_5\leqslant 90\\
& x_1,x_2,x_3,x_4,x_5\geqslant 0
\end{aligned}$$

(提示: 基本解只有一个变量组成.)

6. 下表表示一个特殊的单纯形迭代. 所有变量是非负的. 不论是极大化问题还是极小化问题, 这个表都不是最优的. 因此, 当一个非基变量进入到解中, 它可能会增加 z 或减少 z, 或者保持不变, 这依赖于进基的非基变量的参数.

基	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	解
z	0	−5	0	4	−1	−10	0	0	620
x_8	0	3	0	−2	−3	−1	5	1	12
x_3	0	1	1	3	1	0	3	0	6
x_1	1	−1	0	0	6	−4	0	0	0

(a) 将变量分为基变量和非基变量, 并给出所有变量的当前值.

*(b) 假定问题是极大化类型, 确定可能改进 z 值的非基变量. 如果每个这样的变量进入基本解, 确定相应的离基变量 (如果有的话) 和相应 z 值的改变量. 要求不使用高斯–若尔当行运算.

(c) 重复 (b), 假定问题是极小化类型.

(d) 当选择非基变量进入解时, 哪个 (哪些) 非基变量将不引起 z 值的变化?

7. 考虑图 3.6 的二维解空间.

(a) 假设目标函数已知如下:

$$\max\ z = 3x_1 + 6x_2$$

如果单纯形迭代从点 A 开始, 确定到最优点 E 的路径.

(b) 确定进基变量、可行性条件的相应的比值以及 z 值的改变量, 假定初始迭代在点 A 开始, 目标函数已知如下:

$$\max\ z = 4x_1 + x_2$$

(c) 重复 (b), 假定目标函数是

$$\max\ z = x_1 + 4x_2$$

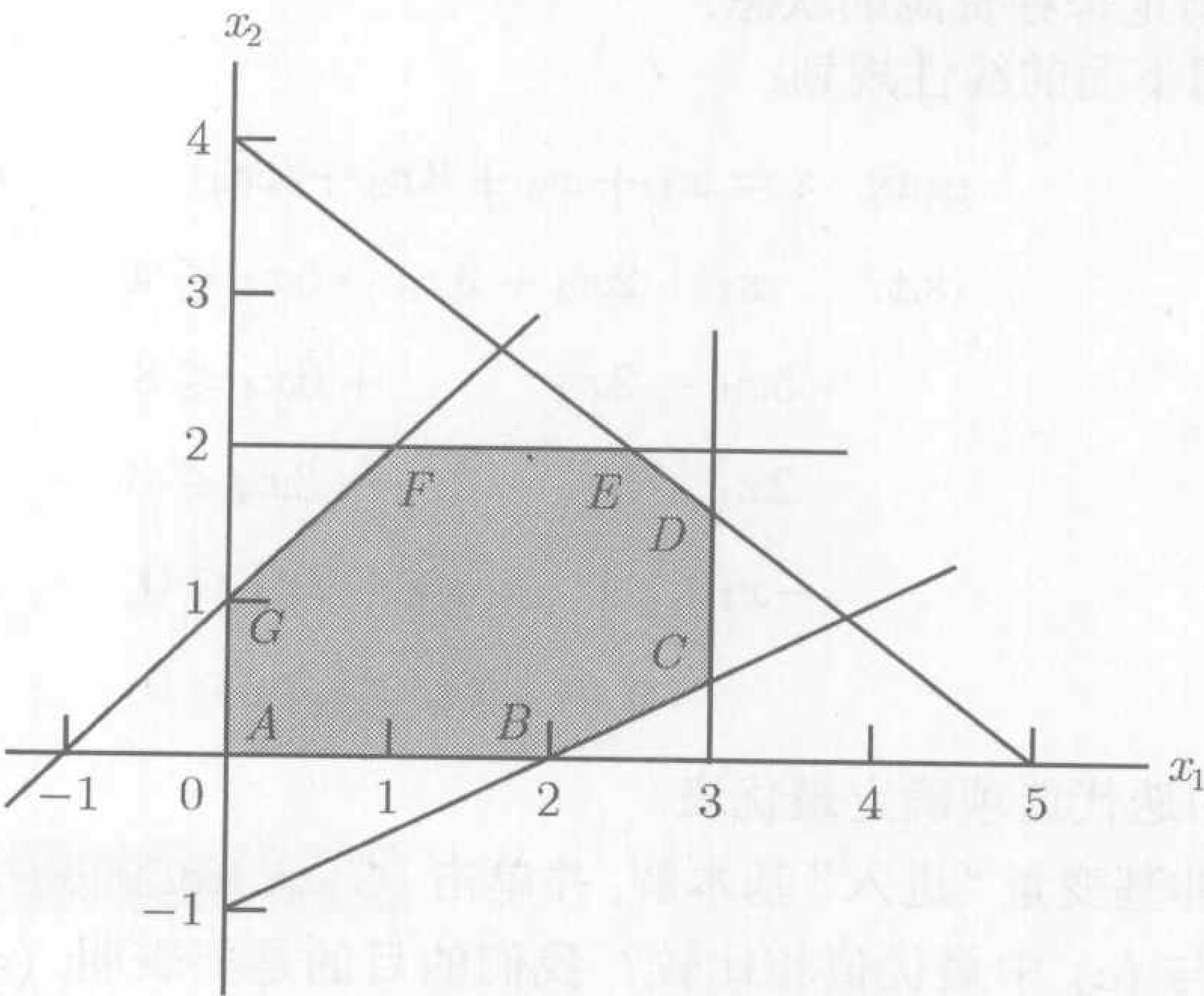


图 3.6 习题 3.3B 第 7 题的解空间

8. 考虑下面的线性规划:

$$
\begin{aligned}
\max \quad & z = 16x_1 + 15x_2 \\
\text{s.t.} \quad & 40x_1 + 31x_2 \leqslant 124 \\
& -x_1 + \quad x_2 \leqslant 1 \\
& x_1 \qquad\quad \leqslant 3 \\
& x_1, x_2 \geqslant 0
\end{aligned}
$$

(a) 用单纯形法求解此问题, 这里进基变量是 z 行中绝对值最大的负系数所对应的非基变量.

(b) 用单纯形算法重解此问题, 始终选择 z 行中绝对值最小的负系数所对应的非基变量作为进基变量.

(c) 比较 (a) 与 (b) 的迭代次数. 选择进基变量作为最负 z 行系数的非基变量是否得到较少的迭代次数? 关于最优性条件, 我们可以得出什么样的结论?

(d) 假定最优化问题变成 -1 乘上 z 求极小. 这个变化如何影响单纯形迭代?

***9.** 在例 3.3-1 中, 说明如何从最优表中确定次最优的 z 值.

10. 你能够扩充第 9 题中的方法, 求出第三最优的 z 值吗?

11. Gutchi 公司制造钱包、梳妆袋和背包. 生产需要皮革和人工合成材料, 且皮革是稀有原料. 生产过程需要两种类型的技能工人: 缝纫与修整. 下表给出了生产 3 种产品所使用资源的可用量和产品每单位的利润.

资源	单位资源需求			日可用量
	钱包	梳妆袋	背包	
皮革 (英尺^2)	2	1	3	42 英尺^2
缝纫 (小时)	2	1	2	40 小时
修整 (小时)	1	0.5	1	45 小时
销售价格 (美元)	24	22	45	

(a) 为该问题建立一个线性规划, 并求最优解 (用 TORA、Excel 规划求解或AMPL).

(b) 从最优解中确定每种资源的状态.

12. TORA 实验. 考虑下面的线性规划:

$$
\begin{aligned}
\max \quad & z = x_1 + x_2 + 3x_3 + 2x_4 \\
\text{s.t.} \quad & x_1 + 2x_2 - 3x_3 + 5x_4 \leqslant 4 \\
& 5x_1 - 2x_2 \qquad\quad + 6x_4 \leqslant 8 \\
& 2x_1 + 3x_2 - 2x_3 + 3x_4 \leqslant 3 \\
& -x_1 \qquad\quad + x_3 + 2x_4 \leqslant 0 \\
& x_1,\ x_2,\ x_3, x_4 \geqslant 0
\end{aligned}
$$

(a) 用 TORA 的迭代选项确定最优表.

(b) 选择任意的非基变量“进入”基本解, 并单击 Next Iteration 产生相应的迭代. 新的目标值如何与 (a) 中最优值相比较? 我们的目的是要表明, (a) 中的表是最优的, 因为没有一个非基变量能够改进目标值.

13. TORA 实验. 在第 12 题中, 用 TORA 求次最好的最优解.

3.4 人工初始解

正如例 3.3-1 所演示的那样, 所有约束是 (⩽) 并且有非负右端项的线性规划方便地提供了全部为松弛变量的初始基本可行解. 包含 (=) 和/或 (⩾) 约束的模型就不是如此.

带有 (=) 和 (⩾) 约束初始“坏状态”线性规划的求解过程是使用在最初迭代中扮演松弛变量角色的**人工变量** (artificial variable), 然后, 在稍后的迭代中恰当地处理它们. 这里介绍两个密切相关的方法: 大 M 方法和两阶段法.

3.4.1 大 M 方法

大 M 方法以等式形式的线性规划开始 (见 3.1 节). 如果第 i 个等式约束没有松弛变量 (或能够扮演松弛变量角色的变量), 那么将人工变量 R_i 加入到初始解中, 类似于所有松弛变量为基本解的情况. 然而, 人工变量并不是原始线性规划模型的一部分, 对于这些变量, 在目标函数中对它们指定非常高的**惩罚** (penalty), 强迫它们在最优解中 (最终) 等于零. 如果问题有可行解, 这种情况总会发生. 下面的规则表明了如何在极大化和极小化的情况下指定惩罚:

人工变量的惩罚规则

已知 M 为一个充分大的正数 (用数学的语言描述, 即 $M \to \infty$), 人工变量的目标系数表示成适当的惩罚, 如果

$$\text{人工变量的目标系数} = \begin{cases} -M, & \text{在极大化问题中} \\ M, & \text{在极小化问题中} \end{cases}$$

例 3.4-1

$$\begin{aligned} \min z &= 4x_1 + x_2 \\ \text{s.t. } 3x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &\geqslant 6 \\ x_1 + 2x_2 &\leqslant 4 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$

用 x_3 作为第 2 个约束的剩余变量, 用 x_4 作为第 3 个约束的松弛变量, 问题的

等式形式如下：

$$
\begin{aligned}
\min\ z &= 4x_1 + x_2 \\
\text{s.t.}\ 3x_1 + x_2 &= 3 \\
4x_1 + 3x_2 - x_3 &= 6 \\
x_1 + 2x_2 + x_4 &= 4 \\
x_1, x_2, x_3, x_4 &\geqslant 0
\end{aligned}
$$

第 3 个方程有松弛变量 x_4，但第 1 个和第 2 个方程却没有. 因此，我们在前两个方程中分别增加人工变量 R_1 和 R_2，并在目标函数中用 $MR_1 + MR_2$ 惩罚它们 (因为我们是在求极小). 导出的线性规划如下：

$$
\begin{aligned}
\min\ z &= 4x_1 + x_2 + MR_1 + MR_2 \\
\text{s.t.}\ 3x_1 + x_2 + R_1 &= 3 \\
4x_1 + 3x_2 - x_3 + R_2 &= 6 \\
x_1 + 2x_2 + x_4 &= 4 \\
x_1, x_2, x_3, x_4, R_1, R_2 &\geqslant 0
\end{aligned}
$$

相应的初始基本解由 $(R_1, R_2, x_4) = (3, 6, 4)$ 给出.

从在计算机上求解问题的观点来看，必须假定 M 是一个数值. 然而，在实际所有的教科书 (包括本书的前 7 个版本) 中，M 在所有的单纯形表中都用作代数运算. 其结果是增加了没有必要的困难层面，这个层面可以用 M 的适当取值而简化避免 (使用计算机时，任何时候都可以这样做). 在本版中，我们将放弃沿用 M 作代数运算的传统，并用数值替代它. 当然我们的意图是，在不丢失问题本质的情况下，简化表达.

我们使用的 M 值应该有多大? 答案依赖于初始线性规划的数据. 注意，*M 相对于初始目标系数*必须是充分大，以使得它将起到惩罚作用迫使人工变量在最优解中取值为零. 与此同时，因为计算机是求解线性规划的主要工具，我们不能希望 M 太大 (即使数学上它应该趋于无穷)，因为当非常大的值与非常小的数值在一起运算时，这可能导致潜在的、严重的舍入误差. 在当前的例子中，x_1 和 x_2 的目标系数分别是 4 和 1. 因此，似乎合理的设置是 $M = 100$.

使用 $M = 100$，初始单纯形表给出如下 (为了方便起见，去掉 z 列，因为它在所有的迭代中不变)：

基	x_1	x_2	x_3	R_1	R_2	x_4	解
z	−4	−1	0	−100	−100	0	0
R_1	3	1	0	1	0	0	3
R_2	4	3	−1	0	1	0	6
x_4	1	2	0	0	0	1	4

在继续进行单纯形方法计算之前，我们需要将 z 行与表的其余部分保持一致. 针对本例，在表中，$x_1 = x_2 = x_3 = 0$，产生初始基本解 $R_1 = 3, R_2 = 6, x_4 = 4$. 这个解产生 $z = 100 \times 3 + 100 \times 6 = 900$(而不是 z 行右端项当前的显示的 0 值). 这种不一致是由于 R_1 和 R_2 在 z 行有非零的系数 $(-100, -100)$ 这一事实造成的 (与例 3.3-1 中所有松弛变量为初始解相比较，那里 z 行松弛变量的系数是零).

我们能够消除这个矛盾，在 z 行选用适当的约束方程替换出 R_1 和 R_2. 针对本例，注意到在 R_1 行和 R_2 行突出显示的元素 $(= 1)$. 用 100 乘上 R_1 行和 R_2 行的每一行，求和之后再加到 z 行上，这样就在目标行中替换出 R_1 和 R_2，也就是，

$$\text{新的 } z \text{ 行} = \text{旧的 } z \text{ 行} + (100 \times R_1 \text{ 行} + 100 \times R_2 \text{ 行})$$

因此，修改表变成 (请验证!)

基	x_1	x_2	x_3	R_1	R_2	x_4	解
z	696	399	-100	0	0	0	900
R_1	3	1	0	1	0	0	3
R_2	4	3	-1	0	1	0	6
x_4	1	2	0	0	0	1	4

注意，$z = 900$，现在与初始基本可行解 $R_1 = 3, R_2 = 6, x_4 = 4$ 是一致的.

最后这张表使我们可以应用单纯形法并使用单纯形的最优性和可行性条件了，正如我们在 3.3.2 节所做的一样. 因为我们是求目标函数的极小值，在 z 行有最正系数 $(= 696)$ 的变量 x_1 进入基本解. 可行性条件的最小比指定 R_1 作为离基变量 (请验证!)

一旦进基变量和离基变量确定后，用所熟知的高斯–若尔当运算就可以计算出新的表.

基	x_1	x_2	x_3	R_1	R_2	x_4	解
z	0	167	-100	-232	0	0	204
x_1	1	$\frac{1}{3}$	0	$\frac{1}{3}$	0	0	1
R_2	0	$\frac{5}{3}$	-1	$-\frac{4}{3}$	1	0	2
x_4	0	$\frac{5}{3}$	0	$-\frac{1}{3}$	0	1	3

最后的单纯表显示 x_2 和 R_2 分别是进基变量与离基变量. 连续使用单纯形计算，再需要两步的迭代即可达到最优解，最优值为 $x_1 = \frac{2}{5}, x_2 = \frac{9}{5}, z = \frac{17}{5}$ (用 TORA 验证!)

注意，人工变量 R_1 和 R_2 在第 1 次和第 2 次迭代时离开基本解，其结果与在目标函数中惩罚它们的概念是相一致的.

评注 如果线性规划没有可行解 (也就是说，约束是不相容的)，惩罚项 M 的使用

将不能迫使在最终的单纯形迭代中人工变量取零值. 在这种情况下, 最终的单纯形表将至少包括一个人工变量取正值. 3.5.4 节将解释这种情形.

习题 3.4A

1. 用手工计算, 完成例 3.4-1 的单纯形迭代, 并得到最优解.
2. TORA 实验. 用 TORA 的 Iterations ⇒ M-method 模块 (文件 toraEx3.4-1.txt) 产生例 3.4-1 的单纯形迭代. 比较 $M=1$、$M=10$ 和 $M=1\,000$ 对求解的影响. 从这个实验中能引出什么结论?
3. 在例 3.4-1 中, 对于下列每一种 (独立的) 情况确定初始表, 在替换出所有人工变量后作出相应的 z 行:

   *(a) 第 3 个约束是 $x_1+2x_2\geqslant 4$　(b) 第 2 个约束是 $4x_1+3x_2\leqslant 6$

   (c) 第 2 个约束是 $4x_1+3x_2=6$　(d) 目标函数是 $\max\ z=4x_1+x_2$
4. 考虑下列这组约束:

$$-2x_1+3x_2=3 \quad (1)$$
$$4x_1+5x_2\geqslant 10 \quad (2)$$
$$x_1+2x_2\leqslant 5 \quad (3)$$
$$6x_1+7x_2\leqslant 3 \quad (4)$$
$$4x_1+8x_2\geqslant 5 \quad (5)$$
$$x_1,x_2\geqslant 0$$

   对于下列每个问题, 在替换出所有人工变量后作出 z 行:

   (a) $\max\ z=5x_1+6x_2$, 约束为 (1), (3), (4).

   (b) $\max\ z=2x_1-7x_2$, 约束为 (1), (2), (4), (5).

   (c) $\min\ z=3x_1+6x_2$, 约束为 (3), (4), (5).

   (d) $\min\ z=4x_1+6x_2$, 约束为 (1), (2), (5).

   (e) $\min\ z=3x_1+2x_2$, 约束为 (1) 和 (5).
5. 考虑下列约束集:

$$x_1+x_2+x_3=7$$
$$2x_1-5x_2+x_3\geqslant 10$$
$$x_1,x_2,x_3\geqslant 0$$

   求解下列目标函数对应的每一个问题:

   (a) $\max\ z=2x_1+3x_2-5x_3$　(b) $\min\ z=2x_1+3x_2-5x_3$

   (c) $\max\ z=x_1+2x_2+x_3$　(d) $\min\ z=4x_1-8x_2+3x_3$

*6. 考虑问题:

$$\max\ z=2x_1+4x_2+4x_3-3x_4$$
$$\text{s.t.}\quad x_1+x_2+x_3=4$$
$$x_1+4x_2+x_4=8$$
$$x_1,\ x_2,\ x_3,\ x_4\geqslant 0$$

   这个问题表明, 两个等式约束中的 x_3 和 x_4 能够担任松弛变量的角色. 它们不同于松弛变量是因为它们在目标函数中有非零的系数. 我们能够用 x_3 和 x_4 作为初始变量, 但是, 像

在人工变量的情况一样, 在执行单纯形迭代之前, 目标函数中这两个初始变量必须被替换出去. 用 x_3 和 x_4 作为初始基变量而不用任何人工变量求解此问题.

7. 用 x_3 和 x_4 作为初始基本可行变量求解下面的问题. 像在第 6 题一样, 不用任何人工变量.

$$\begin{aligned} \min\ z &= 3x_1 + 2x_2 + 3x_3 \\ \text{s.t.}\quad x_1 + 4x_2 + x_3 & \geqslant 7 \\ 2x_1 + x_2 \qquad + x_4 & \geqslant 10 \\ x_1,\ x_2,\ x_3,\ x_4 & \geqslant 0 \end{aligned}$$

8. 考虑问题:

$$\begin{aligned} \max\ z &= x_1 + 5x_2 + 3x_3 \\ \text{s.t.}\quad x_1 + 2x_2 + x_3 &= 3 \\ 2x_1 - x_2 \qquad &= 4 \\ x_1,\ x_2,\ x_3 & \geqslant 0 \end{aligned}$$

变量 x_3 担任松弛变量的角色. 因此, 第 1 个约束不需要人工变量. 然而, 第 2 个约束需要人工变量. 用这两个变量 (即第 1 个约束中的 x_3 和第 2 个约束中的 R_2) 作初始解来求解这个问题.

9. 说明, 大 M 方法如何指出下面的问题没有可行解.

$$\begin{aligned} \max\ z &= 2x_1 + 5x_2 \\ \text{s.t.}\quad 3x_1 + 2x_2 & \geqslant 6 \\ 2x_1 + x_2 & \leqslant 2 \\ x_1,\ x_2 & \geqslant 0 \end{aligned}$$

3.4.2 两阶段法

在大 M 方法中, 惩罚值 M 的使用, 按定义要求, 相对于实际模型的目标系数 M 而言必须很大, 这可能会导致大的舍入误差, 造成单纯形计算精度降低. 两阶段法采用连同常数 M 一起去掉的办法来减少这个困难. 顾名思义, 两阶段法分两个阶段求解线性规划: 阶段 I 试图求一个初始基本可行解, 找到一个解后, 调用阶段 II 求解原问题.

两阶段法的概括

阶段 I 将问题变成等式约束形式, 并在约束中增加必要的人工变量 (如在大 M 方法一样), 以保证找到一个初始基本解. 接下来, 求相应方程的基本解, 使得无论线性规划是求极大化还是极小化, *总是*使人工变量之和达到最小. 如果其和的最小值为正, 则线性规划问题没有可行解, 此过程结束 (这告诉我们, 正的人工变量表示原约束不满足). 否则, 进行阶段 II.

阶段 II 使用阶段 I 得到的可行解作为*原始*问题的初始基本可行解.

例 3.4-2

我们使用与例 3.4-1 相同的问题.

阶段 I

$$
\begin{aligned}
\min\ & r = R_1 + R_2 \\
\text{s.t.}\ & 3x_1 + x_2 + R_1 = 3 \\
& 4x_1 + 3x_2 - x_3 + R_2 = 6 \\
& x_1 + 2x_2 + x_4 = 4 \\
& x_1, x_2, x_3, x_4, R_1, R_2 \geqslant 0
\end{aligned}
$$

相应的单纯形表如下:

基	x_1	x_2	x_3	R_1	R_2	x_4	解
r	0	0	0	-1	-1	0	0
R_1	3	1	0	1	0	0	3
R_2	4	3	-1	0	1	0	6
x_4	1	2	0	0	0	1	4

像在大 M 方法中一样, 在 r 行中的 R_1 和 R_2 用下列计算替换:

$$
\text{新的 } r \text{ 行} = \text{旧的 } r \text{ 行} + (1 \times R_1 \text{ 行} + 1 \times R_2 \text{ 行})
$$

新的 r 行用于求解问题的阶段 I, 这将产生下面的最优表 (用 TORA 的 Iterations ⇒ Two-phase Method 来验证):

基	x_1	x_2	x_3	R_1	R_2	x_4	解
r	0	0	0	-1	-1	0	0
x_1	1	0	$\frac{1}{5}$	$\frac{3}{5}$	$-\frac{1}{5}$	0	$\frac{3}{5}$
x_2	0	1	$-\frac{3}{5}$	$-\frac{4}{5}$	$\frac{3}{5}$	0	$\frac{6}{5}$
x_4	0	0	1	1	-1	1	1

因为最小值 $r = 0$, 阶段 I 产生了基本可行解 $x_1 = \frac{3}{5}$, $x_2 = \frac{6}{5}$, $x_4 = 1$. 此时, 人工变量已完成了它的使命, 从而我们能够从表中连同它们所在的列一起去掉, 并转入阶段 Ⅱ.

阶段 Ⅱ

在去掉人工变量列后, 原始问题如下:

$$
\begin{aligned}
\min\ & z = 4x_1 + x_2 \\
\text{s.t.}\ & x_1 + \tfrac{1}{5}x_3 = \tfrac{3}{5} \\
& x_2 - \tfrac{3}{5}x_3 = \tfrac{6}{5}
\end{aligned}
$$

$$x_3 + x_4 = 1$$
$$x_1,\ x_2,\ x_3,\ x_4 \geqslant 0$$

本质上, 阶段 I 是转换原始约束方程的一个过程, 如果问题的初始基本可行解存在的话, 阶段 I 就会提供它. 因此, 相应于阶段 Ⅱ 问题的单纯形表给出如下:

基	x_1	x_2	x_3	x_4	解
z	-4	-1	0	0	0
x_1	1	0	$\frac{1}{5}$	0	$\frac{3}{5}$
x_2	0	1	$-\frac{3}{5}$	0	$\frac{6}{5}$
x_4	0	0	1	1	1

还有, 因为基变量 x_1 和 x_2 在 z 行有非零的系数, 使用下列计算将非零系数替换出去:

$$\text{新的 } z \text{ 行} = \text{旧的 } z \text{ 行} + (4 \times x_1 \text{ 行} + 1 \times x_2 \text{ 行})$$

因此, 阶段 Ⅱ 的初始单纯形表如下:

基	x_1	x_2	x_3	x_4	解
z	0	0	$\frac{1}{5}$	0	$\frac{18}{5}$
x_1	1	0	$\frac{1}{5}$	0	$\frac{3}{5}$
x_2	0	1	$-\frac{3}{5}$	0	$\frac{6}{5}$
x_4	0	0	1	1	1

因为我们在求极小, x_3 必须进基. 应用单纯形法经过一步迭代即可产生最优解 (用 TORA 验证).

评注 实际上, 几乎所有的商业软件包都使用两阶段法来求解线性规划. 实际工作中, 大 M 方法由于有潜在不利的舍入误差而可能从来不用. 将大 M 方法包含于本书中纯粹是由于历史的原因, 因为它的发展先于两阶段法的发展.

人工变量和它所在的列在阶段 I 的最后一张表被去掉, 这只有在它们全是*非基变量*时才会发生 (正如例 3.4-2 所解释的那样). 如果在阶段 I 的最后一张表中有一个或多个人工变量是*基变量*(取零值), 则必须采用下面的附加步, 才能在阶段 Ⅱ 开始之前将人工变量去掉.

第 1 步 选择一个零人工变量离开基本解, 并指定它所在的行作为*枢轴行*. 进基变量可以是枢轴行具有系数*非零* (正的或负的) 的*任意*非基 (非人工) 变量. 完成相应的单纯形法迭代.

第 2 步 从表中去掉 (刚刚离基的) 人工变量的列. 如果所有的零人工变量已被去掉, 转入阶段 Ⅱ. 否则, 转回第 1 步.

第 1 步的目的是保持基变量的可行性将不受影响, 不论其枢轴元素是正的还

是负的, 总能使一个零人工变量变为非基变量. 在习题 3.4B 中的第 5 题和第 6 题说明了这种情形. 第 7 题提供了进一步关于阶段 I 计算的细节.

习题 3.4B

***1.** 在阶段 I 中, 如果线性规划是极大化类型, 试解释, 我们在阶段 I 中为什么不对人工变量的和求极大.

2. 对于习题 3.4A 中的第 4 题, 写出对应于阶段 I 的目标函数.

3. 用两阶段法求解习题 3.4A 中的第 5 题.

4. 对于下面的问题写出阶段 I, 而后求解 (为方便起见, 可用 TORA), 证明该问题没有可行解.

$$\max \quad z = 2x_1 + 5x_2$$
$$\text{s.t.} \quad 3x_1 + 2x_2 \geqslant 6$$
$$2x_1 + x_2 \leqslant 2$$
$$x_1,\ x_2 \geqslant 0$$

5. 考虑如下问题:

$$\max \quad z = 2x_1 + 2x_2 + 4x_3$$
$$\text{s.t.} \quad 2x_1 + x_2 + x_3 \leqslant 2$$
$$3x_1 + 4x_2 + 2x_3 \geqslant 8$$
$$x_1,\ x_2,\ x_3 \geqslant 0$$

(a) 说明, 阶段 I 将终止在一个取零值的人工基变量 (为方便起见, 可以用 TORA).

(b) 在开始阶段 II 之前, 去掉零人工变量, 然后完成阶段 II 的迭代.

6. 考虑如下问题:

$$\max \quad z = 3x_1 + 2x_2 + 3x_3$$
$$\text{s.t.} \quad 2x_1 + x_2 + x_3 = 2$$
$$x_1 + 3x_2 + x_3 = 6$$
$$3x_1 + 4x_2 + 2x_3 = 8$$
$$x_1,\ x_2,\ x_3 \geqslant 0$$

(a) 说明, 阶段 I 以两个零人工变量终止在基本解中 (为方便起见, 使用 TORA).

(b) 说明, 在阶段 I 的最后用第 5(b) 题, 两个零人工变量中只有一个可以变成非基变量.

(c) 说明, 不能在 (b) 中变成非基变量的零人工变量所对应的原约束一定是多余的 —— 因此, 它所在行和列能够在阶段 II 的初始表中一起去掉.

***7.** 考虑下面的线性规划:

$$\max \quad z = 3x_1 + 2x_2 + 3x_3$$
$$\text{s.t.} \quad 2x_1 + x_2 + x_3 \leqslant 2$$
$$3x_1 + 4x_2 + 2x_3 \geqslant 8$$
$$x_1,\ x_2,\ x_3 \geqslant 0$$

阶段 I 结束时的最优单纯形表已知如下:

基	x_1	x_2	x_3	x_4	x_5	R	解
z	-5	0	-2	-1	-4	0	0
x_2	2	1	1	0	1	0	2
R	-5	0	-2	-1	-4	1	0

请解释, 为什么非基变量 x_1, x_3, x_4, x_5 在阶段 Ⅱ 结束时永远不能为正值. 因此, 这些列可以在我们开始阶段 Ⅱ 之前被去掉. 实际上, 除去这些变量将问题的约束方程简化为 $x_2 = 2$. 这意味着根本不需要运行阶段 Ⅱ, 因为解空间简化为只有一个点.

8. 考虑线性规划模型:

$$
\begin{aligned}
\min\quad & z = 2x_1 - 4x_2 + 3x_3 \\
\text{s.t.}\quad & 5x_1 - 6x_2 + 2x_3 \geqslant 5 \\
& -x_1 + 3x_2 + 5x_3 \geqslant 8 \\
& 2x_1 + 5x_2 - 4x_3 \leqslant 4 \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}
$$

说明如何修改不等式约束以变成仅需要使用一个人工变量的等式约束 (而不用两个人工变量).

3.5 单纯形方法中的特殊情况

这一节考虑在单纯形法中使用的 4 种特殊的情况.

(1) 退化; (2) 可选择最优解; (3) 无界解; (4) 不存在 (或不可行) 解.

在这些特殊情况的研究中, 我们的兴趣有两个方面: (1) 介绍这些情况的理论解释; (2) 提供相应的实际解释, 说明这些特殊的结果在现实生活中意味着什么.

3.5.1 退化

在单纯形方法可行性条件的应用中, 最小比可能循环出现, 可以随意打破这种循环. 当这种情况发生时, 至少有一个基变量在下一次迭代中变为零, 并称新的解是**退化的** (degenerate).

关于退化解没有什么好令人惊讶的, 除了在理论上有些不方便以外, 这种不方便是产生**循环** (cycling), 这一点我们将在后面讨论. 从实用的观点来看, 此条件揭示了模型至少有一个多余的约束. 我们用一个数值例子, 来更加深入理解退化对实际与理论两方面的影响.

例 3.5-1 (退化的最优解)

$$
\begin{aligned}
\max\quad & z = 3x_1 + 9x_2 \\
\text{s.t.}\quad & x_1 + 4x_2 \leqslant 8 \\
& x_1 + 2x_2 \leqslant 4 \\
& x_1,\ x_2 \geqslant 0
\end{aligned}
$$

已知松弛变量 x_3 和 x_4, 下表给出了问题的单纯形迭代:

迭代	基	x_1	x_2	x_3	x_4	解
0	z	-3	-9	0	0	0
x_2 进基	x_3	1	4	1	0	8
x_3 离基	x_4	1	2	0	1	4
1	z	$-\frac{3}{4}$	0	$\frac{9}{4}$	0	18
x_1 进基	x_2	$\frac{1}{4}$	1	$\frac{1}{4}$	0	2
x_4 离基	x_4	$\frac{1}{2}$	0	$-\frac{1}{2}$	1	0
2	z	0	0	$\frac{3}{2}$	$\frac{3}{2}$	18
(最优值)	x_2	0	1	$\frac{1}{2}$	$-\frac{1}{2}$	2
	x_1	1	0	-1	2	0

在迭代 0 中, x_3 和 x_4 均可以是离基变量, 在迭代 1 构成退化, 因为基变量 x_4 出现零值. 再一次迭代后, 达到最优值.

退化的实际含意是什么呢? 看图 3.7 的图解法. 有 3 条线通过最优点 ($x_1 = 0$, $x_2 = 2$). 因为这是一个二维问题, 其点是*超定的*, 而且这些约束中的一个是多余的[①]. 在实际中, 了解一些资源是多余的, 这一信息在解的执行过程中可能是有价值的. 这个信息还可以引导发现模型构造中的缺陷. 不幸的是, 没有有效的计算技巧直接从表中识别多余的约束.

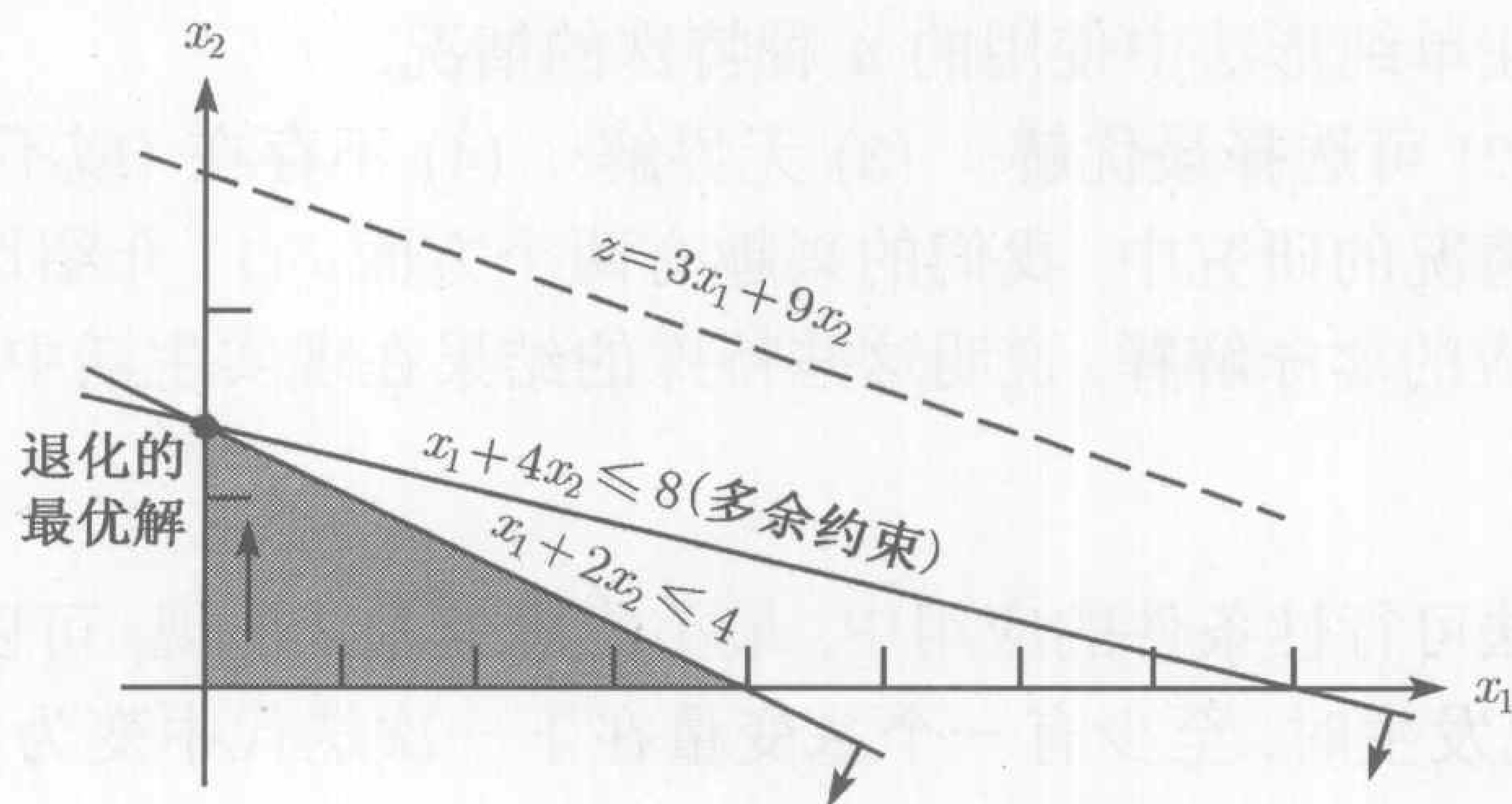


图 3.7 例 3.5-1 中的线性规划退化

从理论的观点来看, 退化有两种含意. 第一个是循环的现象. 看单纯形的迭代 1 和迭代 2, 你将注意到目标值没有改进 ($z = 18$). 因此, 单纯形方法有可能进入一个重复迭代的序列, 永远不改进目标值, 也永远不满足最优性条件 (见习题 3.5A 的第 4 题). 尽管有消除循环的方法, 但这些方法会导致计算速度迅速降低. 针对这个原因, 大多数线性规划程序不包括处理循环的对策, 它依赖于这样一个事实, 即循

① 多余约束一般蕴涵着该约束可以被除去但并不影响可行解空间. 有时会被引用的反例是 $x + y \leqslant 1$, $x \geqslant 1$, $y \geqslant 0$. 这里除去任何一个约束将改变可行空间, 它从一个点变成一个区域. 然而, 即便解空间只由单个的可行点组成, 这也足以说明这个条件是真实的了, 虽然这种情况极不可能出现在现实的线性规划中.

环在实际应用中几乎是不可能出现的.

第二个理论点出现在迭代 1 和迭代 2 的检验中. 两个迭代, 尽管对变量的基与非基的分类不同, 但对于所有的变量和目标值都产生同样的值, 即

$$x_1 = 0, \quad x_2 = 2, \quad x_3 = 0, \quad x_4 = 0, \quad z = 18$$

在迭代 1(退化第一次出现的时候) 是否可以停止计算, 即使它不是最优值? 答案是否定的, 因为解可以像习题 3.5A 中第 2 题演示的那样, 是暂时退化的.

习题 3.5A

*1. 考虑图 3.8 的图形解空间. 假定单纯形迭代在 A 处开始, 最优解出现在点 D 处. 进一步假定目标函数定义如下, 在点 A 处, x_1 首先进入解.

(a) (在图形上) 找出由到达最优点的单纯形法路径所确定的角点.

(b) 确定到达最优点时单纯形迭代所需的最大可能次数, 假定没有循环.

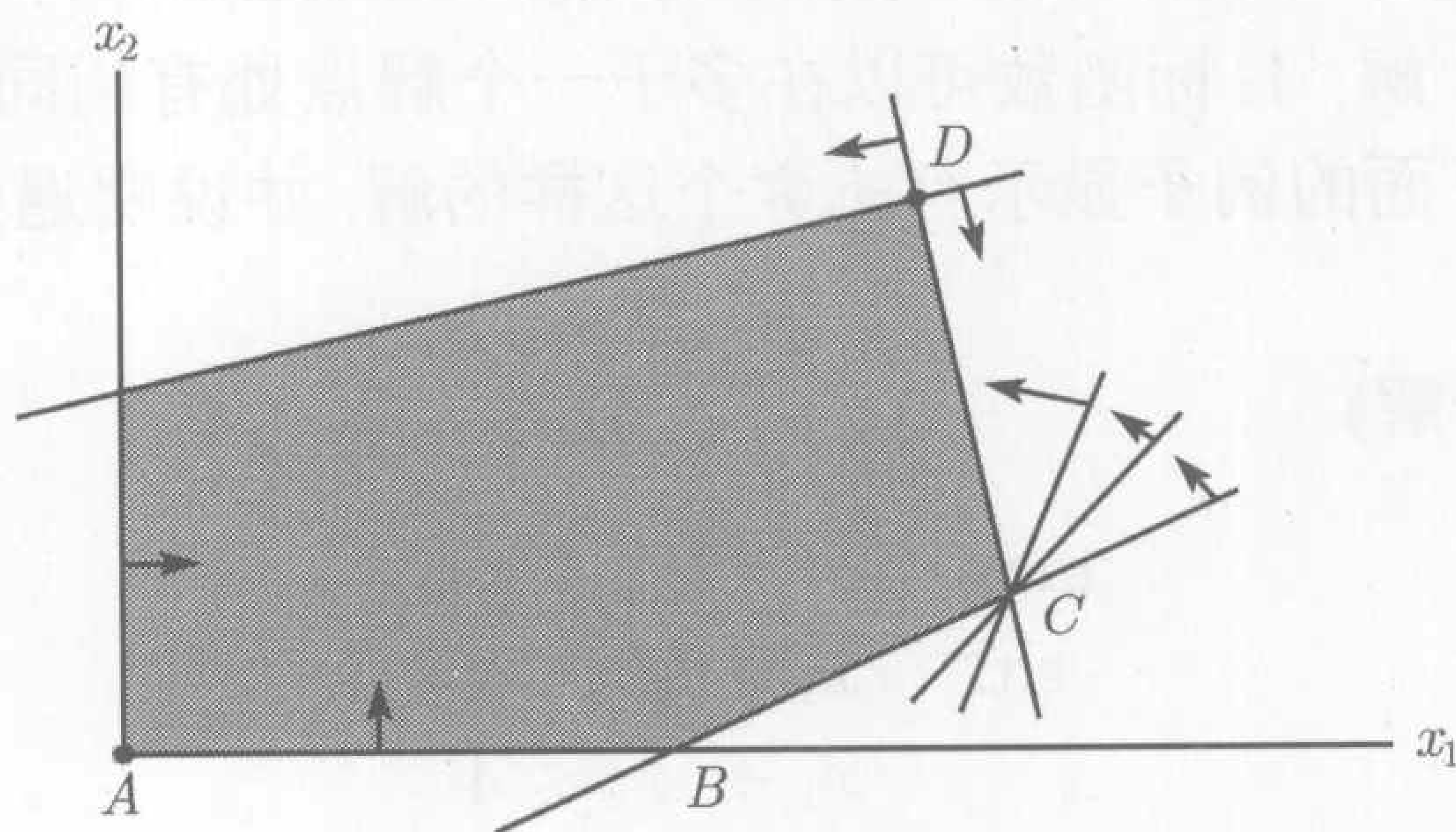


图 3.8 习题 3.5A 第 1 题的解空间

2. 考虑如下线性规划:

$$\begin{aligned} \max \quad & z = 3x_1 + 2x_2 \\ \text{s.t.} \quad & 4x_1 - x_2 \leqslant 8 \\ & 4x_1 + 3x_2 \leqslant 12 \\ & 4x_1 + x_2 \leqslant 8 \\ & x_1, \ x_2 \geqslant 0 \end{aligned}$$

(a) 说明, 相应的单纯形迭代是暂时退化的 (为方便起见, 可以用 TORA).

(b) 用图形方法求解问题来验证这一结果 (这里可以用 TORA 的图形模式).

3. TORA 实验. 考虑第 2 题的线性规划.

(a) 用 TORA 生成单纯形迭代. 达到最优值需要多少次迭代?

(b) 交换约束 (1) 和约束 (3), 并用 TORA 重解该问题. 解此问题需要多少次迭代?

(c) 解释为什么 (a) 和 (b) 的迭代次数是不同的.

4. TORA 实验. 考虑如下线性规划 (由 E. M. Beale 构造的循环例子).

$$\begin{aligned} \max \ & z = \tfrac{3}{4}x_1 - 20x_2 + \tfrac{1}{2}x_3 - 6x_4 \\ \text{s.t.} \ & \tfrac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 \leqslant 0 \end{aligned}$$

$$
\begin{aligned}
\tfrac{1}{2}x_1 - 12x_2 - \tfrac{1}{2}x_3 + 3x_4 &\leqslant 0 \\
x_3 &\leqslant 1 \\
x_1,\ x_2,\ x_3,\ x_4 &\geqslant 0
\end{aligned}
$$

从 TORA 的 SOLVE/MODIFY 菜单中, 选择 Solve ⇒ Algebraic ⇒ Iterations ⇒ All-slack. 接下来用命令 Next iteration "查阅"一系列的单纯形迭代 (不能用 All iterations, 因为单纯形法将产生不确定的循环). 你将注意到, 在迭代 0 处, 初始的所有松弛基本可行解将完全相同地重复出现在迭代 6 中. 这个例子说明了, 在单纯形法的迭代中可能会发生循环, 算法也可能会永远不会收敛到最优解.

有趣的是, 在这个例子中, 如果在这个线性规划的所有系数都乘以适当的倍数而转换成整数值, 循环将不再出现.

3.5.2　可选择最优解

当目标函数平行于非冗余的**紧约束**(binding constraint) (即在最优解处作为方程而被满足的约束) 解, 目标函数可以在多于一个解点处有相同的最优值, 因此产生可选择最优解. 下面的例子显示有*无穷*个这样的解, 并说明遇到这种解的实际意义.

例 3.5-2 (无穷多个解)

$$
\begin{aligned}
\max \quad & z = 2x_1 + 4x_2 \\
\text{s.t.} \quad & x_1 + 2x_2 \leqslant 5 \\
& x_1 + x_2 \leqslant 4 \\
& x_1,\ x_2 \geqslant 0
\end{aligned}
$$

图 3.9 演示了当目标函数平行于某个紧约束时, 可选择最优解如何出现在线性规划模型中. 线段 BC 上任意一点都表示了具有相同目标值 $z = 10$ 的可选择最优解.

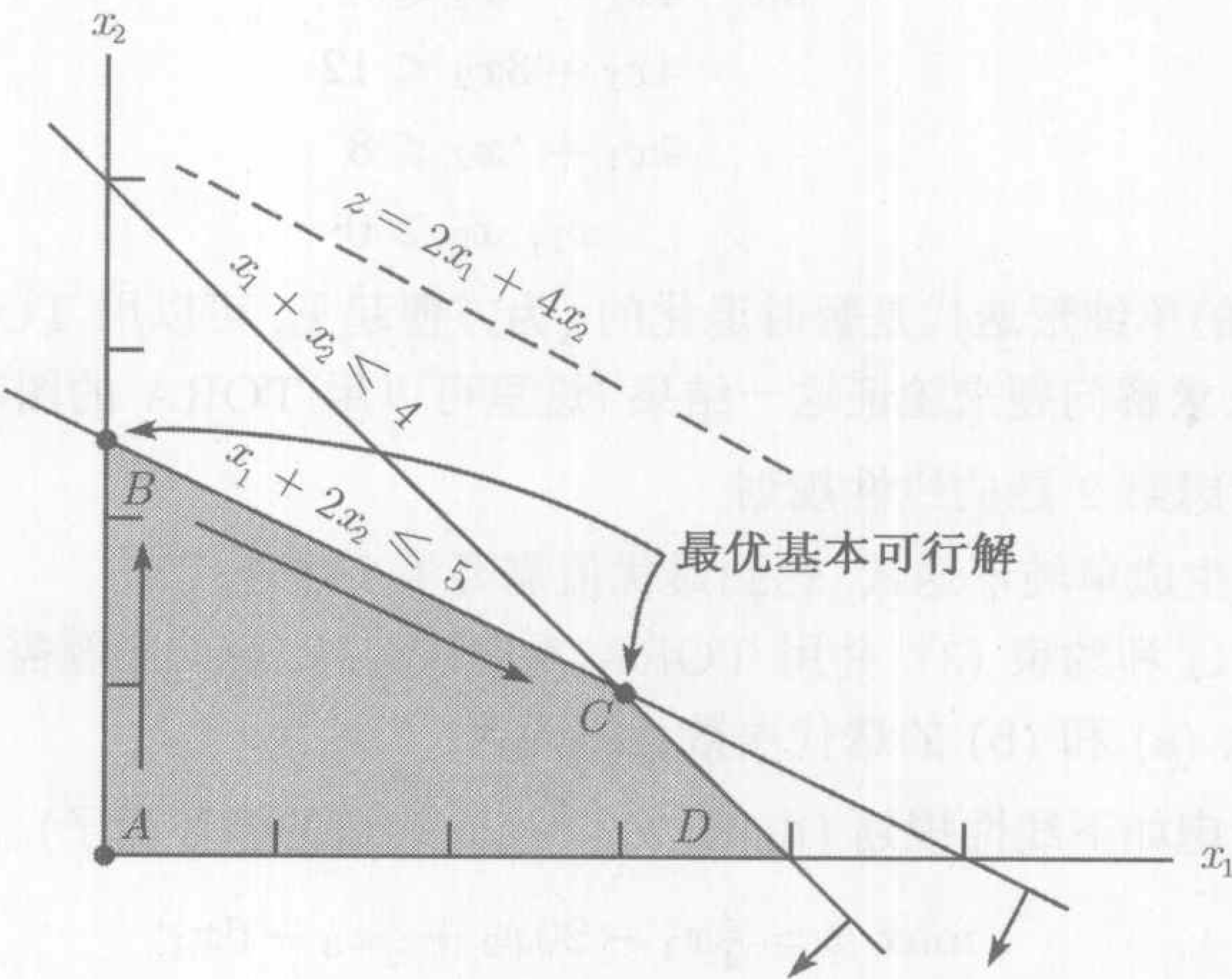


图 3.9　在例 3.5-2 中线性规划的可选择最优解

模型的迭代由下表给出.

迭代	基	x_1	x_2	x_3	x_4	解
0	z	-2	-4	0	0	0
x_2 进基	x_3	1	2	1	0	5
x_3 离基	x_4	1	1	0	1	4
1(最优值)	z	0	0	2	0	10
x_1 进基	x_2	$\frac{1}{2}$	1	$\frac{1}{2}$	0	$\frac{5}{2}$
x_4 离基	x_4	$\frac{1}{2}$	0	$-\frac{1}{2}$	1	$\frac{3}{2}$
2(可选择	z	0	0	2	0	10
最优解)	x_2	0	1	1	-1	1
	x_1	1	0	-1	2	3

迭代 1 给出了最优解 $x_1=0, x_2=\frac{5}{2}$ 和 $z=10$, 它与图 3.9 的点 B 相一致. 我们如何从这个表中知道可选择最优解存在呢? 看一看迭代 1 中 z 行方程非基变量的系数. 非基变量 x_1 的系数是零, 表明 x_1 可以进入基本解而不会改变 z 值, 但会引起变量值的改变. 迭代 2 正是这种情况 —— 令 x_1 进入基本解并迫使 x_4 离基. 新的解点出现在 $C(x_1=3, x_2=1, z=10)$ 处. (TORA 的 Iterations 选项允许你一次确定一个可选择最优解.)

单纯形方法仅能确定两个角点 B 和 C. 从数学角度来说, 我们能够确定线段 BC 上所有的点 (x_1,x_2) 作为点 B 和 C 的非负的加权平均. 因此, 已知

$$B: \quad x_1=0, \quad x_2=\frac{5}{2}$$

$$C: \quad x_1=3, \quad x_2=1$$

则线段 BC 上所有的点由如下式子给出

$$\left.\begin{aligned}\hat{x}_1&=\alpha(0)+(1-\alpha)(3)=3-3\alpha\\ \hat{x}_2&=\alpha(\tfrac{5}{2})+(1-\alpha)(1)=1+\tfrac{3}{2}\alpha\end{aligned}\right\},\quad 0\leqslant\alpha\leqslant 1$$

当 $\alpha=0$, $(\hat{x}_1,\hat{x}_2)=(3,1)$, 它是点 C. 当 $\alpha=1$, $(\hat{x}_1,\hat{x}_2)=(0,\frac{5}{2})$, 它是点 B. 对于 α 在 0 到 1 之间, $(\hat{x}_1,\hat{x}_2)$ 位于 B 到 C 之间.

评注 在实践中, 可选择最优解是有用的, 因为我们可以从许多解中选择而不会损害目标值. 例如, 在前面的例子中, 在 B 处显示, 只有第 2 项活动是正值; 而在 C 处, 两项活动均取正值. 如果例子描述的一种产品混合情形, 生产两种产品比生产一种产品对于满足市场竞争或许更有优势. 在这种情况下, 在 C 处的解可能更吸引人.

习题 3.5B

*1. 对于如下线性规划, 找出 3 个可选择最优基本解, 然后写出由这 3 个基本解所组成的所有非基本可选择最优解的一般表达式.

$$\begin{aligned}
\max \quad & z = x_1 + 2x_2 + 3x_3 \\
\text{s.t.} \quad & x_1 + 2x_2 + 3x_3 \leqslant 10 \\
& x_1 + x_2 \leqslant 5 \\
& x_1 \leqslant 1 \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}$$

注：尽管问题有多于 3 个的可选择基本最优解, 你只需找出其中的 3 个. 为方便起见, 可以使用 TORA.

2. 求解如下线性规划:

$$\begin{aligned}
\max \quad & z = 2x_1 - x_2 + 3x_3 \\
\text{s.t.} \quad & x_1 - x_2 + 5x_3 \leqslant 10 \\
& 2x_1 - x_2 + 3x_3 \leqslant 40 \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}$$

根据最优单纯形表, 证明所有可选择最优解并不是角点 (即是非基本解). 给出这种解空间的一个二维图形的解空间与目标函数的典型示例. (为方便起见, 可以用 TORA).

3. 对于如下线性规划, 说明最优解是退化的, 并且没有一个可选择解是角点 (为方便起见, 可以用 TORA).

$$\begin{aligned}
\max \quad & z = 3x_1 + x_2 \\
\text{s.t.} \quad & x_1 + 2x_2 \leqslant 5 \\
& x_1 + x_2 - x_3 \leqslant 2 \\
& 7x_1 + 3x_2 - 5x_3 \leqslant 20 \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}$$

3.5.3 无界解

在一些线性规划中, 可以无限制地增加变量的值但不破坏任何一个约束, 这意味着解空间至少有一个变量是*无界的*. 其结果是, 目标值可以无限制地增加 (求极大的情况) 或减少 (求极小的情况). 在这种情况下, 解空间和最优目标值都是无界的.

出现无界点可能是由于模型构造得不合理. 在此类模型中最大可能的缺陷是一个或多个非多余约束没有考虑在内, 一些约束的参数 (常数) 可能没有得到正确的估计.

下面的例子说明, 如何能够在单纯形表中发现解空间和目标值均上无界的情况.

例 3.5-3 (无界目标值)

$$\begin{aligned}
\max \quad & z = 2x_1 + x_2 \\
\text{s.t.} \quad & x_1 - x_2 \leqslant 10 \\
& 2x_1 \leqslant 40 \\
& x_1,\ x_2 \geqslant 0
\end{aligned}$$

初始迭代

基	x_1	x_2	x_3	x_4	解
z	-2	-1	0	0	0
x_3	1	-1	1	0	10
x_4	2	0	0	1	40

在初始表中, x_1 和 x_2 均有负的 z 行系数. 因此, 任一个变量都能够改进解. 因为 x_1 有最负的系数, 通常选择它为进基变量. 然而, 在 x_2 下方所有约束的系数 (也就是可行性条件比的分母) 是负的或零. 这意味着没有离基变量, 而且 x_2 能够无限制地增加而不破坏任何约束 (与图 3.5 中最小比的图形解释相比较). 因为 x_2 每增加一个单位, z 将增加 1, x_2 增加至无穷将导致 z 增加至无穷. 因此, 问题无有界解. 这个结果可以在图 3.10 中看到. 在 x_2 方向, 解空间是无界的, 并且 z 值能够无限制地增加.

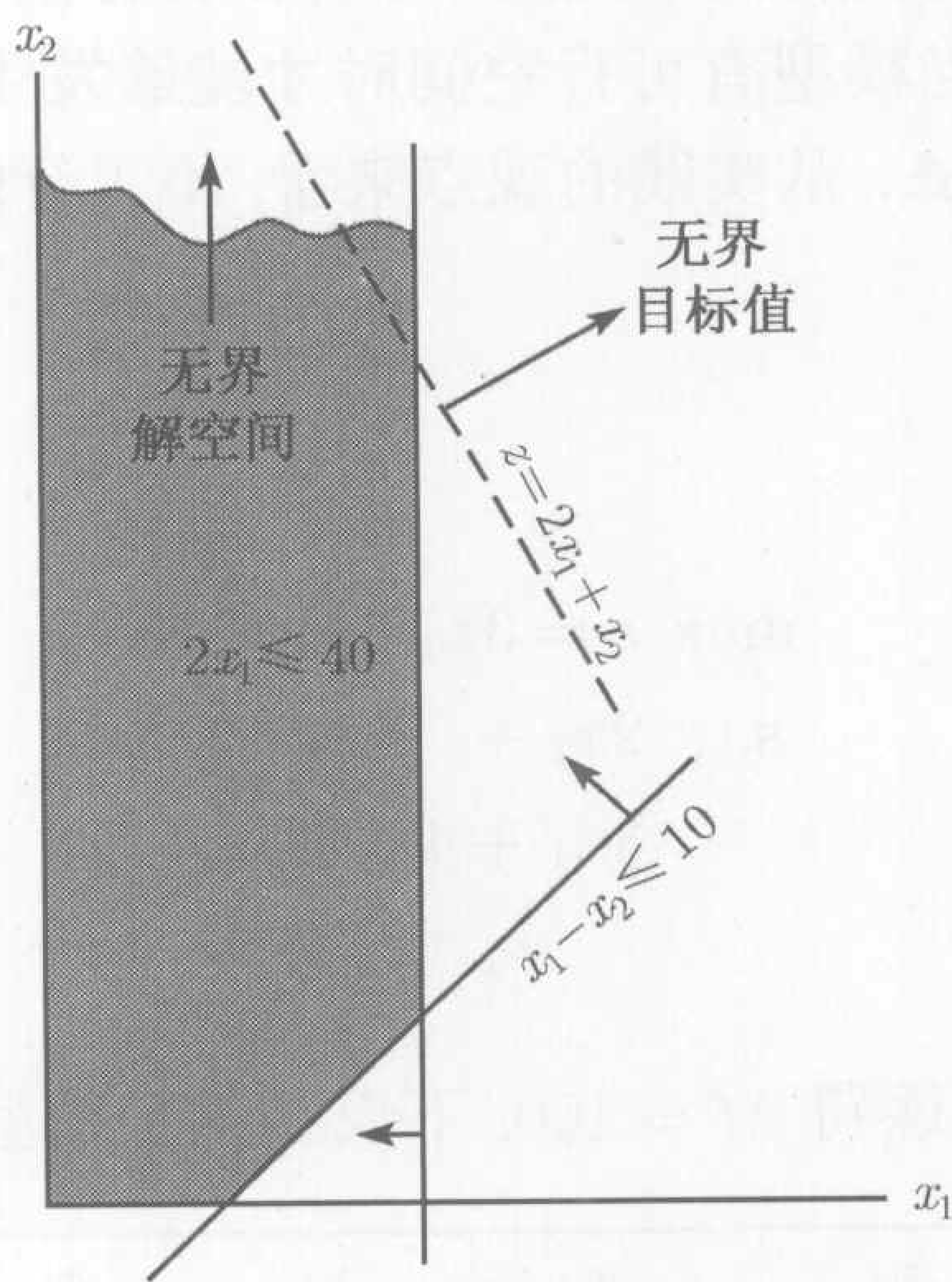


图 3.10 例 3.5-3 中线性规划的无界解

评注 如果我们严格地应用最优性条件让 x_1 进基, 将发生什么? 回答是, 其随后的单纯形表后将最终导致进基变量与 x_2 具有相同的特性, 见习题 3.5C 的第 1 题.

习题 3.5C

1. TORA 实验. 用 TORA 的 Iterations 选项求解例 3.5-3, 并说明尽管解以 x_1 作为进基变量开始 (由最优性条件), 单纯形算法将最终指向一个无界解.

***2.** 考虑线性规划:

$$\begin{aligned} \max\quad & z = 20x_1 + 10x_2 + x_3 \\ \text{s.t.}\quad & 3x_1 - 3x_2 + 5x_3 \leqslant 50 \end{aligned}$$

$$
\begin{aligned}
x_1 \qquad + \ x_3 &\leqslant 10 \\
x_1 - \ x_2 + 4x_3 &\leqslant 20 \\
x_1,\ x_2,\ x_3 &\geqslant 0
\end{aligned}
$$

(a) 观察约束, 确定解空间无界的方向 (x_1、x_2 或 x_3).

(b) 不作进一步的计算, 你能够得到关于最优目标值的结论是什么?

3. 在一些病态构造的线性规划模型, 即使问题可以有一个有界的目标值, 解空间也可以是无界的. 这种情况的发生仅指明模型的构造有缺陷. 在大型问题中, 靠观察发现无界可能是困难的. 设计一种确定解空间是否无界的方法.

3.5.4 不可行解

具有不相容约束的线性规划模型没有可行解. 假如所有 的约束都是 $\leqslant$ 类型并具有非负的右端项, 则这种情况将永远不会出现, 因为松弛变量提供了一个可行解. 对于其他类型的约束, 我们使用人工变量. 尽管在目标函数中惩罚人工变量迫使它们在最优值处取零, 这仅当模型有可行空间时才能够发生. 否则, 至少有一个人工变量在最优迭代中取值为正. 从实践的观点来看, 不可行解空间表明模型的构建有可能是不正确的.

例 3.5-4 (不可行解空间)

考虑如下线性规划:

$$
\begin{aligned}
\max\ z &= 3x_1 + 2x_2 \\
\text{s.t.}\ \ 2x_1 + \ x_2 &\leqslant 2 \\
3x_1 + 4x_2 &\geqslant 12 \\
x_1,\ x_2 &\geqslant 0
\end{aligned}
$$

对于人工变量 R 使用惩罚 $M = 100$, 下表提供了模型的单纯形迭代.

迭代	基	x_1	x_2	x_4	x_3	R	解
0	z	−303	−402	100	0	0	−1 200
x_2 进基	x_3	2	1	0	1	0	2
x_3 离基	R	3	4	−1	0	1	12
1	z	501	0	100	402	0	−396
(伪最优解)	x_2	2	1	0	1	0	2
	R	−5	0	−1	−4	1	4

最优迭代 1 显示人工变量 R 取值为正$(=4)$, 它表明问题不可行. 图 3.11 演示了不可行解空间. 由于允许人工变量是正的, 单纯形方法实际上已经将不等式的方向颠倒, 从 $3x_1 + 4x_2 \geqslant 12$ 变为 $3x_1 + 4x_2 \leqslant 12$(你将如何解释?). 这个结果我们称为**伪最优** (pseudo-optimal) 解.

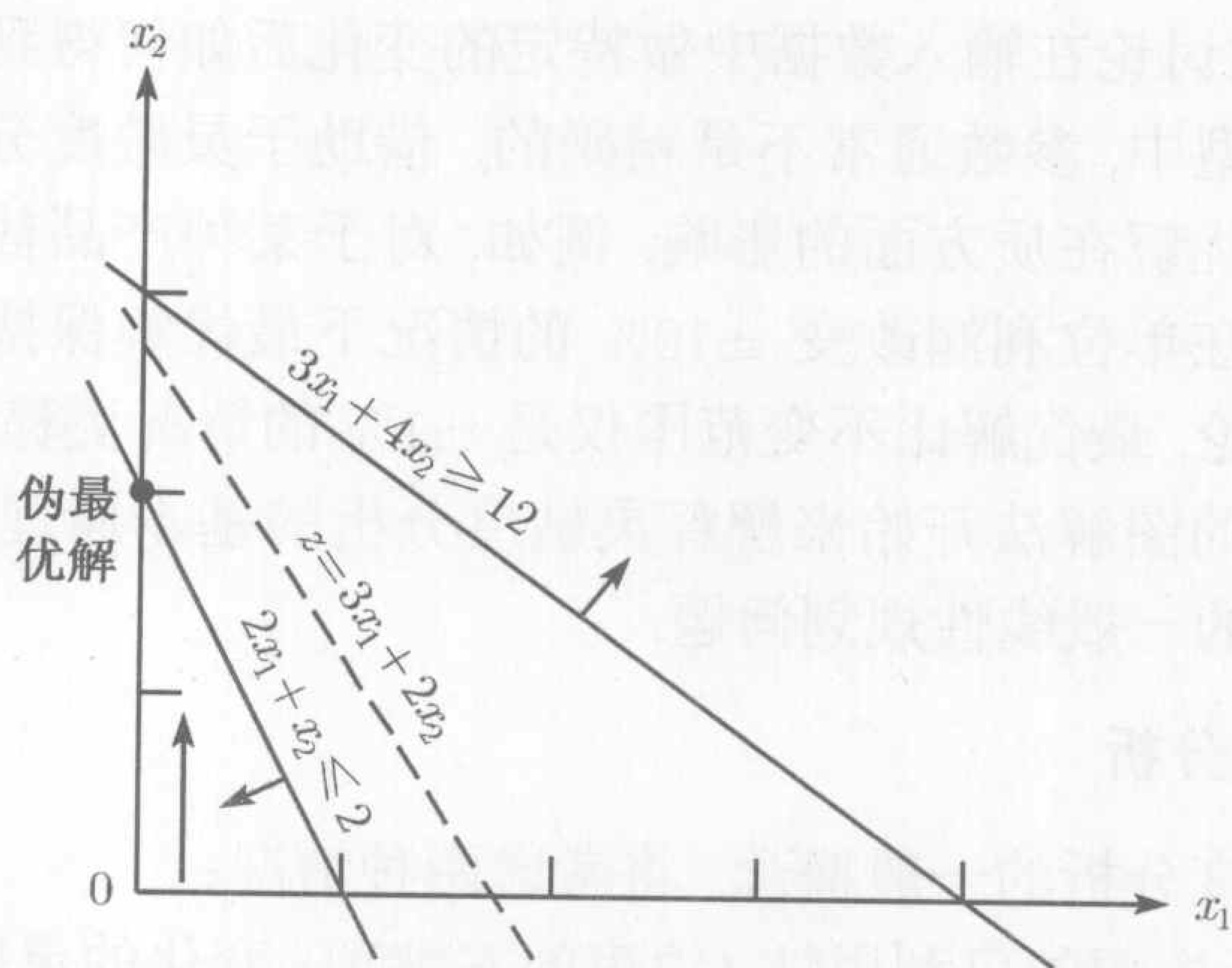


图 3.11 例 3.5-4 的不可行解

习题 3.5D

*1. Toolco 公司生产 3 种类型的工具 T1, T2, T3. 按照下表的数据使用两种原料 M1 和 M2 生产这些工具:

原料	每件工具需要原料的单位数		
	T1	T2	T3
M1	3	5	6
M2	5	3	4

原料 M1 和 M2 的日可利用量分别是 1 000 个单位和 1 200 个单位. 市场部门通知生产经理, 根据他们的研究, 对这 3 种工具的日需求量至少是 500 个单位. 生产部门能够满足需要吗? 如果不能, Toolco 将最多能够提供这 3 种工具各多少?

2. TORA 实验. 考虑如下线性规划模型:

$$\max\ z = 3x_1 + 2x_2 + 3x_3$$

$$\text{s.t.}\ \ 2x_1 + x_2 + x_3 \leqslant 2$$

$$3x_1 + 4x_2 + 2x_3 \geqslant 8$$

$$x_1,\ x_2,\ x_3 \geqslant 0$$

用 TORA 的 Iterations ⇒ M-Method 来说明最优解中包含一个人工基变量, 但取零值. 问题有可行的最优解吗?

3.6 灵敏度分析

在线性规划中, 模型的参数 (输入数据) 能够在一定的限度内变化而不引起最优解的改变. 这些内容涉及*灵敏度分析*, 而这正是这一节的主题. 后面的第 4 章将

研究后最优分析, 它讨论在输入数据中做特定的变化后如何得到新的最优解.

在线性规划模型中, 参数通常不是精确的. 借助于灵敏度分析, 我们能够探索这种不确定性对最优解在质方面的影响. 例如, 对于某种产品估计的单位利润, 如果灵敏度分析揭示在单位利润改变 ±10% 的情况下最优解保持不变, 那么我们就可以得到这样的结论, 最优解比不变范围仅是 ±1% 的情况更稳健.

我们以更具体的图解法开始来解释灵敏度分析的基本原理. 然后可以把它扩充到使用单纯形表的一般线性规划问题.

3.6.1 图形灵敏度分析

这节演示灵敏度分析的一般概念. 将考虑两种情况:

(1) 最优解对于资源的可利用性 (约束的右端项) 变化的灵敏度分析.

(2) 最优解对于单位利润或单位费用 (目标函数的系数) 变化的灵敏度分析.

以二维变量图解线性规划问题作为例子, 来分别考虑上述两种情况.

例 3.6-1 (右端项的变化)

JOBCO 公司在两台机器上生产两种产品. 1 个单位的产品 1 需要 2 小时机器 1 和 1 小时机器 2; 对于产品 2, 1 个单位需要 1 小时机器 1 和 3 小时机器 2. 每个单位产品 1 和产品 2 的收益分别是 30 美元和 20 美元. 每台机器总的日可加工时间是 8 小时.

令 x_1 和 x_2 分别表示产品 1 和产品 2 的日产量, 则线性规划模型给出如下:

$$\begin{aligned} \max\ z &= 30x_1 + 20x_2 \\ \text{s.t.}\ \ 2x_1 + x_2 &\leqslant 8 \quad (\text{机器 } 1) \\ x_1 + 3x_2 &\leqslant 8 \quad (\text{机器 } 2) \\ x_1,\ x_2 &\geqslant 0 \end{aligned}$$

图 3.12 解释了当机器 1 的工作能力发生改变时最优解的变化. 如果日工作能力从 8 小时增加到 9 小时, 新的最优解将在点 G 出现. 把机器 1 的工作能力由 8 小时改变到 9 小时, 所得到的最优值 z 的变化率可由如下方法计算:

$$\begin{pmatrix} \text{机器 1 增加 1} \\ \text{小时工作能力产} \\ \text{生收益的变化率} \\ (\text{点 } C \text{ 变到 } G) \end{pmatrix} = \frac{z_G - z_C}{\text{工作能力的改变}} = \frac{142-128}{9-8} = 14.00(\text{美元/小时})$$

计算出的变化率提供了模型的输入 (资源) 和它的输出 (总收益) 的直接联系, 表示成**资源的单位价值** (unit worth of a resource)(美元/小时), 即资源可用性 (机器能力) 的单位变化引起最优目标函数值的变化. 这意味着机器 1 的能力增加 (减少)1 个单位将增加 (减少) 收益 14 美元. 尽管资源的单位价值是目标函数变化率

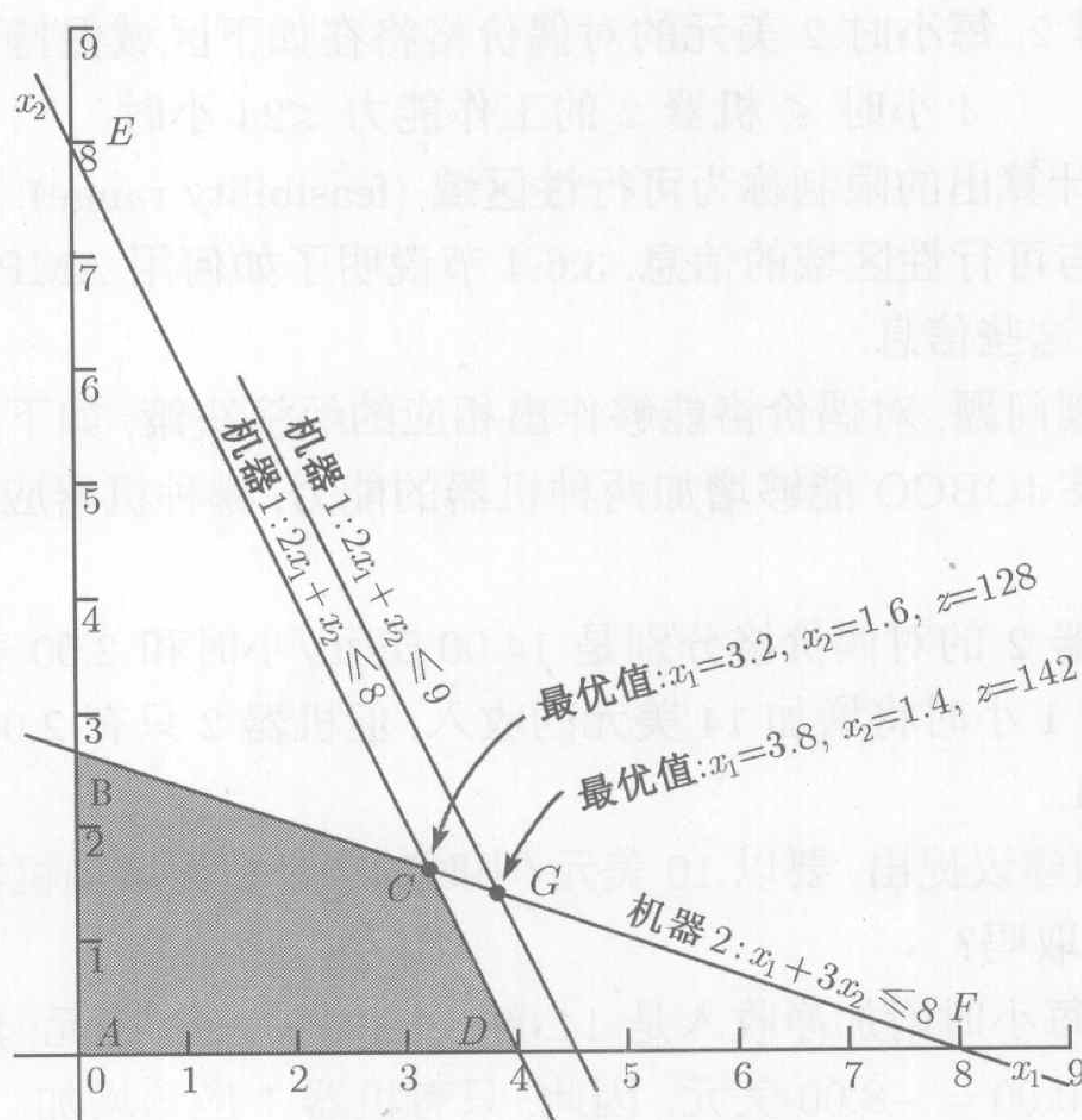


图 3.12 最优解对于改变资源可用量 (约束的右端项) 的图形灵敏度

的恰当描述, 然而它在技术上的名称是**对偶价格** (dual price) 或**影子价格** (shadow price), 它们是现有线性规划文献和所有软件包中的标准用语, 因此, 本书也将沿用这些名称.

从图 3.12 中可以看出, 当机器 1 工作能力变化 (增加或减少), 即其约束平行移动到线段 BF 的任何一点时, 每小时 14 美元的对偶价格仍然有效. 这意味着, 可以按下式计算出给定对偶价格的适用范围:

$$\text{机器 1 工作能力的最小值 [在点 } B=(0,2.67)] \quad = \quad 2\times 0+1\times 2.67=2.67\text{小时}$$

$$\text{机器 1 工作能力的最大值 [在点 } F=(8,0)] \quad = \quad 2\times 8+1\times 0=16\text{小时}$$

因此, 我们能得到如下结论, 每小时 14 美元的对偶价格将在如下区域保持不变:

$$2.67 \text{ 小时} \leqslant \text{机器 1 的工作能力} \leqslant 16 \text{ 小时}$$

在这个区域外变化将产生不同的对偶价格 (每单位价值).

用类似的计算, 可以验证, 机器 2 工作能力的对偶价格是每小时 2 美元, 它在如下区域内变化 (增加或减少) 时保持不变, 即约束平行移动到线段 DE 的任意一点时, 这产生下面的限制区域:

$$\text{机器 2 工作能力的最小值 [在点 } D=(4,0)]=1\times 4+3\times 0=4\text{小时}$$

$$\text{机器 2 工作能力的最大值 [在点 } E=(0,8)]=1\times 0+3\times 8=24\text{小时}$$

结论是, 对于机器 2, 每小时 2 美元的对偶价格将在如下区域保持可适用:

$$4 \text{ 小时} \leqslant \text{机器 2 的工作能力} \leqslant 24 \text{ 小时}$$

机器 1 和机器 2 计算出的限制称为**可行性区域** (feasibility range). 所有软件包均提供关于对偶价格与可行性区域的信息. 3.6.4 节说明了如何用 AMPL、Excel 规划求解和 TORA 生成这些信息.

关于线性规划问题, 对偶价格能够作出相应的经济决策, 如下列问题所示:

问题 1 如果 JOBCO 能够增加两种机器的能力, 哪种机器应该有更高的优先权?

机器 1 和机器 2 的对偶价格分别是 14.00 美元/小时和 2.00 美元/小时. 这意味着, 机器 1 增加 1 小时将增加 14 美元的收入, 但机器 2 只有 2.00 美元. 因此, 优先权应该是机器 1.

问题 2 一项建议提出, 要以 10 美元/小时的额外费用增加机器 1 和机器 2 的能力, 这项建议可取吗?

对于机器 1, 每小时附加净收入是 $14.00 - 10.00 = 4.00$ 美元; 然而对于机器 2, 净收入是 $2.00 - 10.00 = -8.00$ 美元. 因此, 只有机器 1 应该增加.

问题 3 如果机器 1 的工作能力从现有的 8 小时增加到 13 小时, 这项增加将如何影响最优收益?

对于机器 1 的对偶价格是 14.00 美元, 并在 $[2.67, 16]$ 小时的区域内适用. 所提出的增加到 13 小时落在可行区域内. 因此, 收入的增加量是 $14.00(13 - 8) = 70$ 美元, 这意味着总收入将增加到 (当前收益 + 收益的变化)$= 128 + 70 = 198$ 美元.

问题 4 假定机器 1 的工作能力可增加到 20 小时, 这项增加将如何影响最优收益?

所提出的这项改变已超出了 $[2.67, 16]$ 小时这个保持对偶价格 14.00 美元可适用的区域. 因此, 我们只能立刻做出增加到 16 小时的结论. 超过部分需要进一步的计算才能得出答案 (见第 4 章). 记住, 落在可行性区域之外*并不*意味着问题没有解. 它只意味着我们没有充分的信息*立刻*作出决策.

问题 5 我们知道, 只要资源的改变在可行性区域内, 最优目标值的改变就等于 (对偶价格 × 资源的改变). 相应变量的最优值改变了多少?

变量的最优值将一定会改变. 然而, 我们从图解得到的信息水平并不能充分确定新值. 3.6.2 节将从代数角度介绍灵敏度分析问题, 它将提供这方面的细节.

习题 3.6A

1. 一公司生产两种产品 A 和 B. 单位收入分别是 2 美元和 3 美元. 两种原料 M1 和 M2 用于两种产品的生产中, 日可用量分别是 8 个单位和 18 个单位. 1 个单位的产品 A 用 2 个单位的 M1 和 2 个单位的 M2, 1 个单位的产品 B 用 3 个单位的 M1 和 6 个单位的 M2.

(a) 确定 M1 和 M2 的对偶价格和它们的可行性区域.

(b) 假定在每单位 30 美分的费用下可额外得到 4 个单位的 M1. 你推荐额外购买吗?

(c) 公司为每单位 M2 最多支付多少钱?

(d) 如果 M2 的可用量增加 5 个单位, 确定相应的最优收入.

*2. Wild West 生产两种类型的牛仔帽. 1 型帽需要的劳动时间是 2 型帽的 2 倍. 如果所有的可用劳动时间全部用于生产 2 型帽, 公司一天能生产总共 400 顶. 对于这两种类型的帽子市场限制分别是每天 150 顶和 200 顶. 收入是每顶 1 型帽 8 美元, 2 型帽 5 美元.

(a) 用图解法确定最大收入下的每种类型帽子的数量.

(b) 确定生产能力的对偶价格 (依照 2 型帽计算) 和它的可用区域.

(c) 如果关于 1 型帽的限制减少到 120 顶, 用对偶价格确定最优收入的相应效果.

(d) 2 型帽的市场份额的对偶价格是多少? 按照对每顶帽子计算出的价值, 市场份额能够增加多少?

例 3.6-2 (目标系数的变化)

图 3.13 显示了例 3.6-1 提出的 JOBCO 问题的图解空间. 最优值出现在点 $C(x_1 = 3.2, x_2 = 1.6, z = 128)$ 处. 收入单位数的变化 (即目标函数的系数) 将改变 z 的斜率. 然而, 正如从图形中看到的, 只要目标函数位于直线 BF 和直线 DE 之间, 最优点将保持在点 C, 这 2 个约束确定了最优点. 这意味着存在一个关于目标函数系数的区域, 在这个区域内最优解在 C 处保持不变.

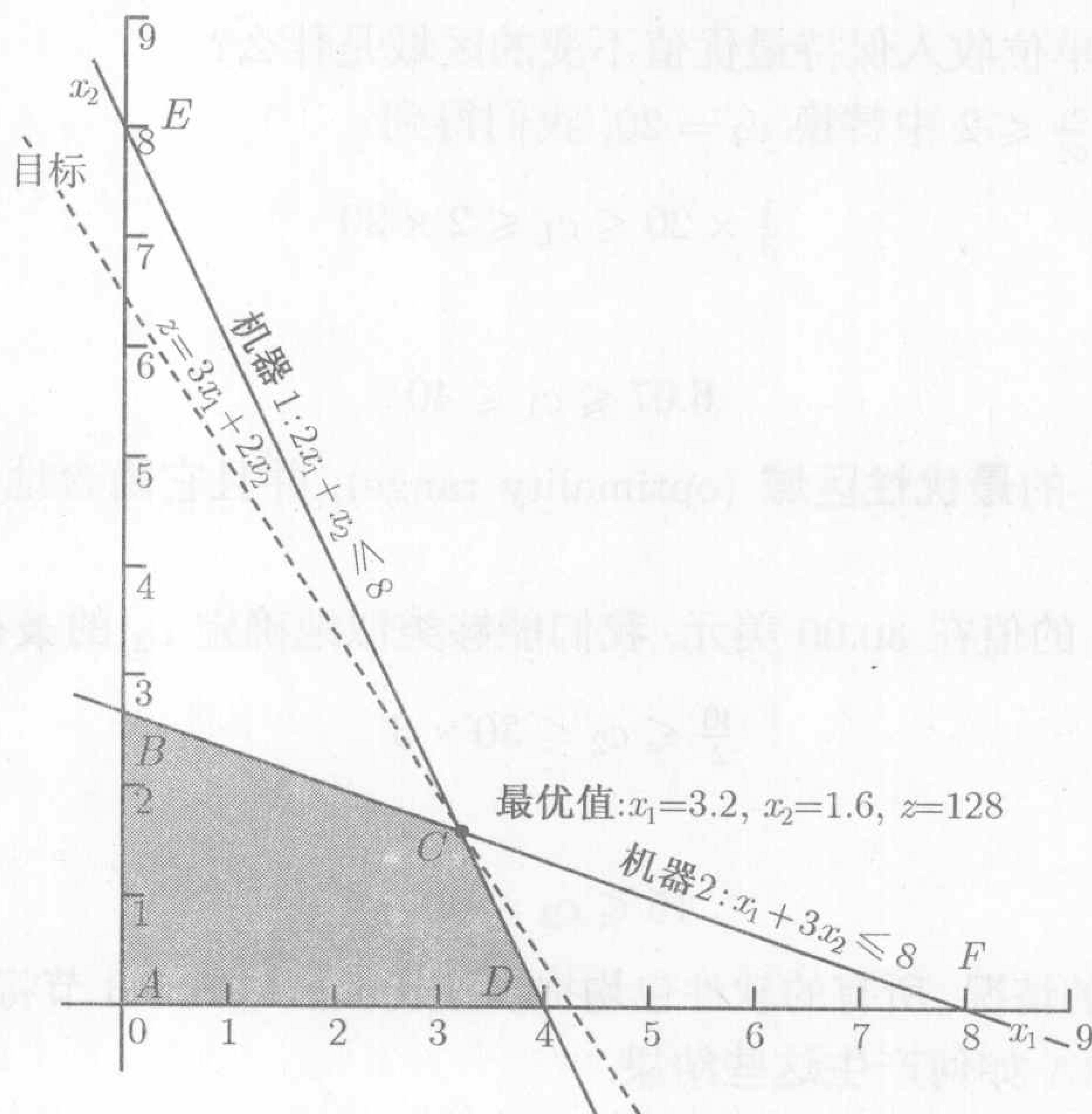


图 3.13 最优解对于改变收入单位数 (目标函数的系数) 的图形灵敏度

可以写出目标函数的一般形式：

$$\max \quad z = c_1x_1 + c_2x_2$$

现在想象, 直线 z 在 C 处转动, 并且它能够沿顺时针方向和逆时针方向旋转. 最优解始终保持在点 C, 只要 $z = c_1x_1 + c_2x_2$ 介于直线 $x_1 + 3x_2 = 8$ 和直线 $2x_1 + x_2 = 8$ 之间. 这意味着比值 $\frac{c_1}{c_2}$ 可以在 $\frac{1}{3}$ 与 $\frac{2}{1}$ 之间变化, 这产生下列条件:

$$\frac{1}{3} \leqslant \frac{c_1}{c_2} \leqslant \frac{2}{1} \quad \text{或} \quad 0.333 \leqslant \frac{c_1}{c_2} \leqslant 2$$

这个信息能够直接提供有关最优解的答案, 如下列问题所示:

问题 1　假设产品 1 和产品 2 的单位收入分别改变到 35 美元和 25 美元. 当前的最优解保持不变吗?

新的目标函数是

$$\max \ z = 35x_1 + 25x_2$$

解在 C 处将保持最优, 因为 $\frac{c_1}{c_2} = \frac{35}{25} = 1.4$ 保持在最优性区域 $[\frac{1}{3}, 2]$ 之内. 当比值落在这个区域之外, 需要增加额外的计算来求出新的最优解 (见第 4 章). 注意, 尽管变量在最优点 C 的值保持不变, z 的最优值变到 $35 \times (3.2) + 25 \times (1.6) = 152.00$ 美元.

问题 2　假定产品 2 的单位收入固定在当前值 $c_2 = 20.00$ 美元上. 相应于 c_1, 也就是产品 1 的单位收入保持最优值不变的区域是什么?

在条件 $\frac{1}{3} \leqslant \frac{c_1}{c_2} \leqslant 2$ 中替换 $c_2 = 20$, 我们得到

$$\tfrac{1}{3} \times 20 \leqslant c_1 \leqslant 2 \times 20$$

或

$$6.67 \leqslant c_1 \leqslant 40$$

这个区域称为 c_1 的**最优性区域** (optimality range), 并且它隐含地假定 c_2 固定在 20.00 美元.

如果固定 c_1 的值在 30.00 美元, 我们能够类似地确定 c_2 的**最优性区域**. 因此,

$$\tfrac{30}{2} \leqslant c_2 \leqslant 30 \times 3$$

或

$$15 \leqslant c_2 \leqslant 90$$

对于右端项的情况, 所有的软件包均提供最优性区域. 3.6.4 节将说明 AMPL、Excel 规划求解和 TORA 如何产生这些结果.

评注　尽管这一节中的资料仅处理二维变量的问题, 但这些结果奠定了对一般线性规划进行灵敏度分析 (这些内容将在 3.6.2 节和 3.6.3 节中介绍) 的基础.

习题 3.6B

1. 考虑习题 3.6A 的第 1 题.

(a) 确定保持最优值不变的 $\frac{c_A}{c_B}$ 的最优性条件.

(b) 确定 c_A 和 c_B 的最优性区域, 假定另一个系数在它当前值保持不变.

(c) 如果 c_A 和 c_B 的单位收入分别同时变到 5 美元和 4 美元. 确定新的最优解.

(d) 如果在 (c) 中, 一次改变一个收入值, 最优解会发生什么变化?

2. 在例 2.2-1 的 Reddy Mikks 模型中:

(a) 确定外墙涂料的单位收入与内墙涂料单位收入比的区域.

(b) 如果每吨外墙涂料的收入保持在每吨 5 000 美元不变, 确定保持当前最优解不变的内墙涂料的最大单位收入.

(c) 如果由于市场原因, 内墙涂料的单位收入必须降到 3 000 美元, 当前最优产品混合会改变吗?

***3.** 在习题 3.6A 的第 2 题中:

(a) 在保持当前最优解不变的情况下, 确定两种类型帽子单位收入比的最优性区域.

(b) 使用由 (a) 得到的信息, 如果两种类型的帽子每单位收入是相同的, 最优解将如何变化?

3.6.2 代数灵敏度分析——右端项的变化

3.6.1 节用图形解确定了对偶价格 (资源的单位价值) 和它们的可行性区域. 本节将把这种分析扩充到一般线性规划模型. 一个数值例子 (TOYCO 模型) 将用于作这方面的介绍.

例 3.6-3 (TOYCO 模型)

TOYCO 通过 3 种操作装配 3 种玩具——玩具火车、玩具卡车和玩具汽车. 对于这 3 种操作可用时间限制分别是每天 430 分钟、460 分钟和 420 分钟, 玩具火车、玩具卡车和玩具汽车的单位收入分别是 3 美元、2 美元和 5 美元. 每辆玩具火车在 3 种操作的装配时间分别是 1 分钟、3 分钟和 1 分钟. 每辆玩具卡车和每辆玩具汽车相应的时间是 (2,0,4) 和 (1,2,0) 分钟 (零时间表示不使用该项操作).

令 x_1, x_2, x_3 分别表示每天装配玩具火车、玩具卡车和玩具汽车的单位数量, 相应的线性规划模型如下:

$$\begin{aligned}
\max\ z &= 3x_1 + 2x_2 + 5x_3 \\
\text{s.t.}\quad & x_1 + 2x_2 + x_3 \leqslant 430 \ (\text{操作 } 1) \\
& 3x_1 + 2x_3 \leqslant 460 \ (\text{操作 } 2) \\
& x_1 + 4x_2 \leqslant 420 \ (\text{操作 } 3) \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}$$

分别用 x_4, x_5, x_6 表示约束操作 1、操作 2 和操作 3 对应的松弛变量, 最优表是

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	4	0	0	1	2	0	1 350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_6	2	0	0	-2	1	1	20

这个结果告诉我们要生产玩具卡车 100 辆和玩具汽车 230 辆, 但不生产玩具火车. 相应的收入是 1 350 美元.

对偶价格的确定 在增加松弛变量 x_4, x_5, x_6 后, 模型的约束可写成:

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 430 \text{ (操作 1)} \\ 3x_1 \qquad + 2x_3 + x_5 &= 460 \text{ (操作 2)} \\ x_1 + 4x_2 \qquad + x_6 &= 420 \text{ (操作 3)} \end{aligned}$$

或

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 430 - x_4 \text{ (操作 1)} \\ 3x_1 \qquad + 2x_3 &= 460 - x_5 \text{ (操作 2)} \\ x_1 + 4x_2 \qquad &= 420 - x_6 \text{ (操作 3)} \end{aligned}$$

借助于这个表达式, 松弛变量具有与操作时间相同的单位 (分钟). 因此, 我们能够说, 在松弛变量上*减少* 1 分钟等价于在操作时间上*增加* 1 分钟.

可以用上述信息从最优表的 z 方程中确定对偶价格:

$$z + 4x_1 + x_4 + 2x_5 + 0x_6 = 1\,350$$

这个方程可以写成:

$$\begin{aligned} z &= 1\,350 - 4x_1 - x_4 - 2x_5 - 0x_6 \\ &= 1\,350 - 4x_1 + 1(-x_4) + 2(-x_5) + 0(-x_6) \end{aligned}$$

已知松弛变量值的*减少*等价于在它的操作时间上的*增加*, 因此我们得到

$$\begin{aligned} z = 1\,350 - 4x_1 &+ 1 \times \text{增加操作 1 的时间} \\ &+ 2 \times \text{增加操作 2 的时间} \\ &+ 0 \times \text{增加操作 3 的时间} \end{aligned}$$

这个方程显示了 (1) 在操作 1 上时间增加 1 分钟 z 将增加 1 美元; (2) 在操作 2 上时间增加 1 分钟 z 将增加 2 美元; (3) 在操作 3 上时间增加 1 分钟 z 将不改变.

概括最优表的 z 行如下:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	4	0	0	1	2	0	1 350

它直接产生对偶价格, 其值由下表所示:

资源	松弛变量	松弛变量的最优 z 方程系数	对偶价格 (美元/分钟)
操作 1	x_4	1	1
操作 2	x_5	2	2
操作 3	x_6	0	0

操作 3 的零对偶价格意味着, 分配给这个操作更多的生产时间没有经济效益. 这个结果有意义, 因为资源已经充裕了, 证据就是, 在最优解中与操作 3 相对应的松弛变量是正值 (= 20). 至于操作 1 和操作 2 的每一个松弛变量, 增加 1 分钟将分别提高收入 1 美元和 2 美元. 对偶价格还表明, 当分配额外的资源时, 操作 2 将得到更高的优先权, 因为它的对偶价格是操作 1 的 2 倍.

上述计算说明对偶价格是如何从最优表的 $\leqslant$ 约束中确定的. 对于 $\geqslant$ 约束, 相同的想法仍然是可适用的, 但是对偶价格将与相应的 $\leqslant$ 约束的符号相反. 至于约束是等式的情况, 从最优单纯形表确定对偶价格需要一些“相关的”计算, 这些将在第 4 章说明.

可行性区域的确定 已确定了对偶价格, 下一步我们来说明如何确定保持它们有效性的可行性区域. 令 D_1, D_2, D_3 分别是分配给操作 1、操作 2 和操作 3 的每天生产时间的改变量 (正的或负的). 模型可写成如下形式:

$$\begin{aligned}
\max\ z &= 3x_1 + 2x_2 + 5x_3 \\
\text{s.t.}\quad x_1 + 2x_2 + x_3 &\leqslant 430 + D_1 \text{ (操作 1)} \\
3x_1 \qquad + 2x_3 &\leqslant 460 + D_2 \text{ (操作 2)} \\
x_1 + 4x_2 \qquad &\leqslant 420 + D_3 \text{ (操作 3)} \\
x_1,\ x_2,\ x_3 &\geqslant 0
\end{aligned}$$

我们将考虑同时发生改变的一般情况. 每次只改变一个变量的特殊情况将从这些结果中得到.

具体方法是用所修正的右端项重新计算最优单纯形表, 然后获得保持解可行的条件, 即最优表的右端项保持非负. 为了说明右端项如何重新计算, 我们以修正解列开始, 即在初始单纯形表中使用新的右端项: $430 + D_1$、$460 + D_2$ 和 $420 + D_3$. 因此, 初始表将呈现如下形式:

基	x_1	x_2	x_3	x_4	x_5	x_6	解			
							右端项	D_1	D_2	D_3
z	-3	-2	-5	0	0	0	0	0	0	0
x_4	1	2	1	1	0	0	430	1	0	0
x_5	3	0	2	0	1	0	460	0	1	0
x_6	1	4	0	0	0	1	420	0	0	1

在 D_1, D_2, D_3 下的列与初始基本列 x_4, x_5, x_6 下的那些列是相同的. 这意味着当我们像在初始模型那样完成相同的单纯形迭代时, 两组中的列也一定是完全相同. 实际上, 新的最优表将变成

基	x_1	x_2	x_3	x_4	x_5	x_6	解			
							右端项	D_1	D_2	D_3
z	4	0	0	1	2	0	1 350	1	2	0
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100	$\frac{1}{2}$	$-\frac{1}{4}$	0
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230	0	$\frac{1}{2}$	0
x_6	2	0	0	-2	1	1	20	-2	1	1

新的最优表提供了如下最优解:

$$z = 1\,350 + D_1 + 2D_2$$
$$x_2 = 100 + \tfrac{1}{2}D_1 - \tfrac{1}{4}D_2$$
$$x_3 = 230 + \tfrac{1}{2}D_2$$
$$x_6 = 20 - 2D_1 + D_2 + D_3$$

有趣的是, 正如前面说明的, 新的 z 值确认了操作 1、操作 2 和操作 3 的对偶价格分别是 1, 2, 0.

只要所有变量非负, 则当前解保持可行, 这就导出下列**可行性条件**:

$$x_2 = 100 + \tfrac{1}{2}D_1 - \tfrac{1}{4}D_2 \geqslant 0$$
$$x_3 = 230 + \tfrac{1}{2}D_2 \geqslant 0$$
$$x_6 = 20 - 2D_1 + D_2 + D_3 \geqslant 0$$

任意同步改变 D_1, D_2, D_3, 只要满足这些不等式, 都将保持解可行. 如果所有的条件都满足, 则可在上述等式中通过直接替换 D_1, D_2, D_3 来找到新的最优解.

为解释这些条件的使用, 假定对于操作 1、操作 2 和操作 3 的可利用生产时间分别是 480 分钟、440 分钟和 410 分钟. 则 $D_1 = 480 - 430 = 50$, $D_2 = 440 - 460 = -20$, $D_3 = 410 - 420 = -10$. 在可行性条件中作替换, 我们得到

$$x_2 = 100 + \tfrac{1}{2}(50) - \tfrac{1}{4}(-20) = 130 > 0 \quad \text{(可行)}$$

$$x_3 = 230 + \tfrac{1}{2}(-20) = 220 > 0 \quad \text{(可行)}$$

$$x_6 = 20 - 2(50) + (-20) + (-10) = -110 < 0 \quad \text{(不可行)}$$

计算表明 $x_6 < 0$, 因此, 当前解并没有保持可行. 需要额外的计算才能得出新的解. 这些计算将在第 4 章中后最优分析中讨论.

另一种方法是, 如果资源的变化使得 $D_1 = -30, D_2 = -12, D_3 = 10$, 则

$$x_2 = 100 + \tfrac{1}{2}(-30) - \tfrac{1}{4}(-12) = 88 > 0 \quad \text{(可行)}$$

$$x_3 = 230 + \tfrac{1}{2}(-12) = 224 > 0 \quad \text{(可行)}$$

$$x_6 = 20 - 2(-30) + (-12) + (10) = 78 > 0 \quad \text{(可行)}$$

新的可行解是 $x_2 = 88, x_3 = 224, x_6 = 78$, 且 $z = 3(0) + 2(88) + 5(224) = 1\,296$ 美元. 注意, 最优目标值还可以用如下的方法计算: $z = 1\,350 + 1(-30) + 2(-12) = 1\,296$ 美元.

给出的条件可以专门用于产生各自的可行性区域, 也就是一次只改变一种资源的变化结果 (正如 3.6.1 节所定义的).

情况 1　把操作 1 的时间从 460 分钟改变到 $(460 + D_1)$ 分钟. 这一变化等价于在同步条件中置 $D_2 = D_3 = 0$, 得到

$$\left.\begin{aligned} x_2 &= 100 + \tfrac{1}{2}D_1 \geqslant 0 \quad \Rightarrow D_1 \geqslant -200 \\ x_3 &= 230 > 0 \\ x_6 &= 20 - 2D_1 \geqslant 0 \quad \Rightarrow D_1 \leqslant 10 \end{aligned}\right\} \Rightarrow -200 \leqslant D_1 \leqslant 10$$

情况 2　把操作 2 的时间从 430 分钟变成 $(430 + D_2)$ 分钟. 这一变化等价于在同步条件中置 $D_1 = D_3 = 0$, 得到

$$\left.\begin{aligned} x_2 &= 100 - \tfrac{1}{4}D_2 \geqslant 0 \quad \Rightarrow D_2 \leqslant 400 \\ x_3 &= 230 + \tfrac{1}{2}D_2 \geqslant 0 \quad \Rightarrow D_2 \geqslant -460 \\ x_6 &= 20 + D_2 \geqslant 0 \quad \Rightarrow D_2 \geqslant -20 \end{aligned}\right\} \Rightarrow -20 \leqslant D_2 \leqslant 400$$

情况 3　把操作 3 的时间从 420 分钟变成 $(420 + D_3)$ 分钟. 这一变化等价于在同步条件中置 $D_1 = D_2 = 0$, 得到

$$\left.\begin{aligned} x_2 &= 100 > 0 \\ x_3 &= 230 > 0 \\ x_6 &= 20 + D_3 \geqslant 0 \end{aligned}\right\} \Rightarrow -20 \leqslant D_3 < \infty$$

对于 TOYCO 模型, 我们现在可以概括对偶价格和可行性区域如下[①]:

① 能够利用的线性规划软件包通常给出这个信息作为标准输出. 实际上, 没有软件提供同步条件的所有情形, 大概因为它们的显示太麻烦了, 特别是对于大型线性规划问题.

资源	对偶价格	可行性区域	资源数量 (分钟)		
			最小值	当前值	最大值
操作 1	1	$-200 \leqslant D_1 \leqslant 10$	230	430	440
操作 2	2	$-20 \leqslant D_2 \leqslant 400$	440	440	860
操作 3	0	$-20 \leqslant D_3 < \infty$	400	420	∞

我们一定要注意, 对于任何保持解可行的同步改变, 即使这种改变已超出了各自的可行性区域, 对偶价格仍然是可适用的. 例如, 改变量 $D_1 = 30$, $D_2 = -12$, $D_3 = 100$ 仍然保持解可行, 尽管 $D_1 = 30$ 超出了可行性区域 $-200 \leqslant D_1 \leqslant 10$, 正如下面计算显示:

$$x_2 = 100 + \tfrac{1}{2}(30) - \tfrac{1}{4}(-12) = 118 > 0 \qquad \text{(可行)}$$

$$x_3 = 230 + \tfrac{1}{2}(-12) = 224 > 0 \qquad \text{(可行)}$$

$$x_6 = 20 - 2(30) + (-12) + (100) = 48 > 0 \qquad \text{(可行)}$$

这意味着对偶价格仍然可适用, 因此我们能够从对偶价格中计算出新的最优目标值 $z = 1\,350 + 1(30) + 2(-12) + 0(100) = 1\,356$ 美元.

对上述结果可总结如下:

(1) 对偶价格保持有效, 只要约束右端项的改变量 $D_i, i = 1, 2, \cdots, m$ 在同步改变时满足所有可行性条件, 或相应的 D_i 在单个发生改变时仍落在可行性区域内.

(2) 如果对偶价格不再有效, 因为同步的可行性条件不满足, 或者因为单个的可行性区域被破坏, 此时, 可以用新的 D_i 值重解该问题, 或者应用第 4 章提出的后最优分析来解决.

习题 3.6C①

1. 在 TOYCO 模型中, 假定改变量 D_1, D_2, D_3 在 3 个操作中同时发生.

(a) 如果操作 1、操作 2 和操作 3 的可用量分别变化到 438 分钟、500 分钟和 410 分钟, 用同步条件来说明当前基解保持可行, 并用最优对偶价格确定最优收入的改变量.

(b) 如果 3 个操作的可用量分别变化到 460 分钟、440 分钟和 380 分钟, 用同步条件说明当前基解变为不可行.

***2.** 考虑 TOYCO 模型.

(a) 对于操作 1, 假定超过它当前每天 430 分钟能力的任何附加时间必须依靠每小时 50 美元的加班来获得. 每小时成本包括劳动力和机器运行费两个方面. 对于操作 1, 使用加班在经济上有利吗?

(b) 假定操作 2 的操作员已同意每天加班工作 2 小时, 其加班费是 45 美元/小时. 还有, 操作自身的成本是 10 美元/小时. 这项活动对于每天的净收入有什么影响?

(c) 操作 3 需要加班时间吗?

① 在这组习题中, 你可以发现, 借助于 TORA 能够非常方便地生成最优单纯形表.

(d) 假定操作 1 的每天可用量增加到 440 分钟. 任何超过当前最大能力所使用的加班时间将花费 40 美元/小时. 确定新的最优解, 包括相应的净收入.

(e) 假定操作 2 的可用量每天减少 15 分钟, 并且在正常时间内操作的每小时成本是 30 美元. 减少操作 2 的可用量是有利的吗?

3. 一公司生产 3 种产品 A, B, C. 产品 A 的销售量至少是 3 种产品总销售量的 50%. 然而, 公司每天的销售不能超过 75 个单位的产品 A. 3 种产品使用同一种原料且最大日可用量是 240 磅. 原料的使用比是每单位产品 A 为 2 磅, 每单位产品 B 为 4 磅, 每单位产品 C 为 3 磅. 产品 A, B, C 的单位价格分别是 20 美元、50 美元和 35 美元.

(a) 为公司确定最优的生产策略.

(b) 确定原料资源的对偶价格和它的可允许范围. 如果可用原料增加 120 磅, 确定最优解, 并用对偶价格确定总收入的改变量.

(c) 用对偶价格确定改变产品 A 的需求 ±10 个单位的效果.

4. 某公司一天工作 10 个小时, 在 3 个相继的过程中生产 3 种产品. 下表总结了问题的数据:

产品	生产每单位产品需要的时间 (分钟)			单位价格 (美元)
	过程 1	过程 2	过程 3	
1	10	6	8	4.50
2	5	8	10	5.00
3	6	9	12	4.00

(a) 确定最优生产策略.

(b) 使用对偶价格, 对这 3 种过程投入时间的可能扩大区分优先次序.

(c) 如果能够分配更多的生产时间, 每个过程所附加的每小时的合理费用是多少?

5. Ozark 社区学院继续教育系每学期共提供 30 门课程. 提供的课程通常有 2 种类型. 一种是实用类型, 如木工、文字处理和汽车维护; 另一种是人文类型, 如历史、音乐和美术. 为了满足社区的需求, 每学期每种类型的课程至少提供 10 门. 继续教育系估计提供实用类型和人文类型课程的每门课程的收入分别大约是 1 500 美元和 1 000 美元.

(a) 为该学院设计一组最优课程.

(b) 证明: 一门附加课程的对偶价格是 1 500 美元, 它与每一门实用类型课程的收入相同. 在提供的附加课程方面, 这一结果意味着什么?

(c) 当保证每门课程对总收入的贡献是 1 500 美元时, 能够提供多少更多的课程?

(d) 当增加一门人文类型课程最小需求时, 确定所产生收入的改变量.

***6.** Show & Sell 能够为它的产品在本地电台、电视或在报纸上作广告. 广告预算不超过每月 10 000 美元. 每分钟电台广告的费用是 15 美元, 每分钟电视广告的费用是 300 美元. 报纸广告的费用是 50 美元. Show & Sell 希望电台广告至少是电视广告的 2 倍. 同时, 推荐在一个月内至少使用 5 次报纸广告和不超过 400 分钟的电台广告. 过去的经验表明, 电视广告的效果是电台广告的 50 倍, 是报纸广告的 10 倍.

(a) 为 3 种媒体确定最优预算分配.

(b) 对电台广告和报纸广告设置限制, 从经济角度看合理吗?

(c) 如果月预算增加 50%, 这个结果会导致整体的广告效果按比例增加吗?

7. Burroughs 服装公司为 Walmark 廉价商店生产男式衬衫和女式上衣. Walmark 将接受由 Burroughs 提供的全部产品. 生产过程包括裁剪、缝纫和包装 3 部分. Burroughs 雇用 25 名工人在裁剪车间, 35 名工人在缝纫车间, 以及 5 名工人在包装车间. 工厂每天工作 8 小时, 一周工作 5 天. 下表给出两种服装的时间需求和每件服装的单位价格:

服装	每件服装所需的时间 (分钟)			价格/件 (美元)
	裁剪	缝纫	包装	
男式衬衫	20	70	12	8
女式上衣	60	60	4	12

(a) 为 Burroughs 确定每周的生产计划.

(b) 依照总收入, 确定每小时裁剪、缝纫和包装的价值.

(c) 如果加班时间可用于裁剪与缝纫上, Burroughs 将付给加班的最大钟点费是多少?

8. ChemLabs 公司使用原料 I 和 Ⅱ 生产两种家用清洁剂 A 和 B. 原料 I 和 Ⅱ 的日可用量分别是 150 单位和 145 单位. 1 个单位的清洁剂 A 要消耗 0.5 个单位原料 I 和 0.6 个单位的原料 Ⅱ, 1 个单位的清洁剂 B 要消耗 0.5 个单位原料 I 和 0.4 个单位的原料 Ⅱ. 每个单位的清洁剂 A 和 B 的价格分别是 8 美元和 10 美元. 清洁剂 A 的日需求量在 30 个到 150 个单位之间, 清洁剂 B 的日需求量在 40 个到 200 个单位之间.

(a) 求 ChemLabs 公司生产清洁剂 A 和 B 的最优数量.

(b) 用对偶价格确定, 放松产品 A 和 B 的哪一个需求限制能够改善获利能力.

(c) 如果更多的原料能够以每单位 20 美元获得, 这项做法是可取的吗? 试解释.

(d) 有人建议, 增加 25% 的原料 Ⅱ 以移除生产瓶颈, 这项做法是可取的吗? 试解释.

9. 一条装配线由 3 个相邻的工作站组成, 生产两种型号的收音机 DiGi-1 和 DiGi-2. 下表提供了 3 个工作站的装配时间.

工作站	每件产品所需的时间 (分钟)	
	DiGi-1	DiGi-2
1	6	4
2	5	4
3	4	6

对工作站 1, 2, 3 每天的维护所消耗的时间分别占每个站的总可利用时间的 10%, 14%, 12%, 每个站一天的最大可用时间是 480 分钟.

(a) 公司希望确定最优的产品生产策略, 使得 3 个工作站的空闲 (非使用) 时间最小. 确定工作站的最优利用. [提示: 依照原始变量, 表示成 3 项操作空闲时间 (松弛变量) 的和].

(b) 确定每个工作站减少 1 个百分点日维护时间的价值.

(c) 有人提议, 3 个工作站的运行时间每天增加到 600 分钟, 其附加成本是每分钟 1.50 美元. 这项提议对最优解能有改进吗?

10. Gutchi 公司制造钱包、梳妆袋和背包. 生产需要皮革和人工合成材料, 且皮革是稀有原料. 生产过程需要两种类型的技能劳动：缝纫与修整. 下表给出了生产 3 种产品所使用资源的可用量和 3 种产品的单位价格.

资源	每单位资源需求			日可用量
	钱包	梳妆袋	背包	
皮革 (英尺^2)	2	1	3	42
缝纫 (小时)	2	1	2	40
修整 (小时)	1	0.5	1	45
价格 (美元)	24	22	45	

将问题建立成一个线性规划, 并求最优解. 接下来说明, 下列资源的改变是否保持当前解可行. 对于保持可行性的情况, 确定新的最优解 (变量值和目标函数值).

(a) 可利用皮革增加到 45 英尺^2.
(b) 可利用皮革减少 1 英尺^2.
(c) 可利用缝纫时间变为 38 小时.
(d) 可利用缝纫时间变为 46 小时.
(e) 可利用修整时间减少到 15 小时.
(f) 可利用修整时间增加到 50 小时.
(g) 你会推荐以每小时 15 美元费用雇用额外的缝纫工人吗?

11. HiDec 使用电阻、电容和集成电路片生产两种型号的电子装置. 下表总结了各种情形的数据：

资源	单位资源需求		最大可利用量 (件)
	型号 1(件)	型号 2(件)	
电阻	2	3	1 200
电容	2	1	1 000
集成电路片	0	4	800
单位价格 (美元)	3	4	

令 x_1 和 x_2 分别是型号 1 和型号 2 的生产量. 下面是线性规划模型和相应的最优单纯形表.

$$\max\ z = 3x_1 + 4x_2$$

$$\text{s.t.}\quad 2x_1 + 3x_2 \leqslant 1\,200\ (\text{电阻})$$

$$2x_1 + x_2 \leqslant 1\,000\ (\text{电容})$$

$$4x_2 \leqslant 800\ (\text{集成电路片})$$

$$x_1,\ x_2 \geqslant 0$$

基	x_1	x_2	s_1	s_2	s_3	解
z	0	0	$\frac{5}{4}$	$\frac{1}{4}$	0	1 750
x_1	1	0	$-\frac{1}{4}$	$\frac{3}{4}$	0	450
s_3	0	0	-2	2	1	400
x_2	0	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	100

*(a) 确定每种资源的状态.

*(b) 依照最优收入, 确定电阻、电容和集成电路片的对偶价格.

(c) 对于由 (b) 获得的对偶价格, 确定其可行性区域.

(d) 如果电阻的可用量增加到 1 300 件, 求新的最优解.

*(e) 如果集成电路片的可用量减少到 350 件, 你能够直接从给出的信息中确定新的最优解吗? 试解释.

(f) 如果电容的可用量限制在由 (c) 计算得到的可行性区域内, 确定最优收入的相应区域和由型号 1 与型号 2 生产的单位数量相应的区域.

(g) 新的合同商打算以每个 40 美分的价格销售给 HiDec 额外的电阻, 但仅当 HiDec 至少购买 500 件时方能按照这个价格. HiDec 应该接受这种方案吗?

12. 100% *可行性规则*. 这是一个简化的规则, 它基于约束右端项*各自*的改变量 $D_1, D_2, \cdots, D_m$, 可以用于检验同步的改变是否会保持当前解的可行性. 假定*一次一个*地将约束 i 的右端项 b_i 改变到 $b_i + D_i$, 并且用 3.6.2 节的过程获得相应的可行性区域 $p_i \leqslant D_i \leqslant q_i$. 由定义, 我们有 $p_i \leqslant 0 (q_i \geqslant 0)$, 因为它表示 b_i 最大允许的减少 (增加) 量. 接下来, 如果 D_i 是负的定义 r_i 等于 $\frac{D_i}{p_i}$, 如果 D_i 是正的定义 r_i 等于 $\frac{D_i}{q_i}$. 由定义, 我们有 $0 \leqslant r_i \leqslant 1$. 因此 100% 规则是说, 已知改变量 $D_1, D_2, \cdots, D_m$, 则当前解保持可行的一个*充分* (但不必要) 的条件是 $r_1 + r_2 + \cdots + r_m \leqslant 1$. 如果条件不满足, 则当前解可能保持或者可能不保持可行. 这个规则并不适用于 D_i 落在区域 $[p_i, q_i]$ 之外.

实际上, 100% 规则太弱了, 很多情况下没有什么用. 即使在可行性能够被证实的情况下, 我们还需要用正规的单纯形可行性条件获得新的解. 此外, 3.6.2 节给出的与同步改变量相关的直接计算更直接了当, 也更易于处理.

为显示这一规则的弱点, 我们把它应用到本节习题第 1 题的 (a) 和 (b). 该规则未能证实 (a) 中解的可行性; 也不能应用在 (b) 中, 因为 D_i 的改变量落在可容许区域之外. 第 13 题进一步说明了这一点.

13. 考虑问题

$$
\begin{aligned}
\max\ z &= x_1 + x_2 \\
\text{s.t.}\ 2x_1 + x_2 &\leqslant 6 \\
x_1 + 2x_2 &\leqslant 6 \\
x_1,\ x_2 &\geqslant 0
\end{aligned}
$$

(a) 证明: 最优基本解包含 x_1 和 x_2, 如果一次只考虑改变一个约束, 两个约束的可行性区域分别是 $-3 \leqslant D_1 \leqslant 6$ 和 $-3 \leqslant D_2 \leqslant 6$.

(b) 假定有两种资源, 每种资源以 $\Delta > 0$ 同时增加. 首先, 证明对于所有的 $\Delta > 0$, 基本解始终保持可行. 其次, 证明 100% 规则仅能确认当增加量是在区间 $0 < \Delta \leqslant 3$ 之

间的可行性. 否则, 规则对于区间 $3 < \Delta \leqslant 6$ 不能用, 也不能应用于区间 $\Delta > 6$ 的情形.

3.6.3 代数灵敏度分析——目标函数

3.6.1 节用图形灵敏度分析确定了保持 2 维变量的线性规划解最优性的条件. 本节把这些概念扩充到一般线性规划问题.

简约费用 (reduced cost) 的定义 为便于目标函数灵敏度分析的解释, 首先我们需要定义*简约费用*. 在 TOYCO 模型中 (例 3.6-3), 在最优表中目标 z 行方程是

$$z + 4x_1 + x_4 + 2x_5 = 1\ 350$$

或

$$z = 1\ 350 - 4x_1 - x_4 - 2x_5$$

最优解并不建议玩具火车的生产 ($x_1 = 0$). 这项建议由 z 行方程的信息得到了证实, 因为 x_1 在当前零值的情况下, 每增加一个单位, z 值将下降 4 美元, 即 $z = 1\ 350 - 4 \times (1) - 1 \times (0) - 2 \times (0) = 1\ 346$ 美元.

可以把 z 行方程中 x_1 的系数 (= 4) 作为*费用*的单位, 因为它引起收入 z 的减少. 但这个"费用"来自哪里呢? 我们知道, 在原模型中, x_1 的单位收入是 3 美元. 我们还知道, 每个玩具火车消耗资源 (操作时间), 它本身又导致费用. 因此, 从优化的观点来看, x_1 的"吸引力"依赖于单位收入与一个单位资源消耗的费用的相对值. 这个关系在线性规划文献中被公式化, 定义简约费用如下:

单位简约费用 = 单位消耗资源费用 − 单位收入

为了体会这个定义的意义, 在原始 TOYCO 模型中, 玩具卡车的单位收入 (= 2 美元) 少于玩具火车的单位收入 (= 3 美元). 然而, 最优解选择生产玩具卡车 ($x_2 =$ 100 单位) 而不生产玩具火车 ($x_1 = 0$). 这个 (表面上并不直观) 结果的原因是用于玩具卡车资源 (也就是操作时间) 的单位费用小于它的单位价格. 对于玩具火车却是相反的情况.

借助于给出的*简约费用*的定义, 现在可以看到无利益的变量 (如 x_1) 用下面两种方法使它成为有利可图:

(1) 增加单位收入; (2) 减少所消耗资源的单位费用.

在大多数现实情形中, 单位产品的价格不是一个可以实施的选项, 因为它的值由市场条件决定. 那么现实的选项是减少资源消耗, 比如使生产过程更有效, 这些将在第 4 章讨论.

最优性区域的定义 现在将注意力转向到确定使最优解保持不变的条件. 这部分内容介绍是基于*简约费用*的定义.

在 TOYCO 模型中, 令 d_1, d_2, d_3 分别表示玩具火车、卡车和汽车在单位收入上的改变量. 则目标函数变为

$$\max\ z = (3 + d_1)x_1 + (2 + d_2)x_2 + (5 + d_3)x_3$$

正如 3.6.2 节对于右端项灵敏度分析所做的那样, 我们将首先处理一般情况, 其中目标函数所有的系数同步改变, 然后再用这个结果专门处理一次一个系数发生改变的情况.

由于是同步改变, 在初始表中 z 行呈现为:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	$-3 - d_1$	$-2 - d_2$	$-5 - d_3$	0	0	0	0

当我们在原始模型 (在引进改变量 d_j 之前) 中, 使用相同的进基变量和离基变量序列来产生单纯形表时, 最优迭代将呈现如下形式 (借助于执行单纯形行运算, 可验证情况确实如此)

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	$4 - \frac{1}{4}d_2 + \frac{3}{2}d_3 - d_1$	0	0	$1 + \frac{1}{2}d_2$	$2 - \frac{1}{4}d_2 + \frac{1}{2}d_3$	0	$1\,350 + 100d_2 + 230d_3$
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_6	$-\frac{1}{4}$	0	0	-2	1	1	20

除了简约费用(z 方程系数) 发生改变外, 新的最优表与原始最优表完全相同. 这意味着目标函数系数的改变可以只影响问题的最优性条件.

你确实不需要完成相关的行运算来计算新的简约费用. 新的 z 行的检查说明 d_j 的系数直接来自最优表约束的系数. 计算新的简约费用的一种方便的方法是, 在最优表上增加一个新的顶行和新的最左边列, 正如下面由阴影区域所显示的那样. 在顶行的这些量是相应于每个变量的改变量 d_j. 对于最左边的列, 由 z 行的 1 和每个基变量行中的对应 d_j 构成. 记住, 对于松弛变量 $d_j = 0$.

		d_1	d_2	d_3	0	0	0	
	基	x_1	x_2	x_3	x_4	x_5	x_6	解
1	z	4	0	0	1	2	0	1 350
d_2	x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
d_3	x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
0	x_6	2	0	0	-2	1	1	20

现在, 对于每个变量 (或 z 的值) 计算新的简约费用, 用最左端列相应的元素乘上它所在列的元素, 将它们加在一起, 然后从其和中减去顶行的元素. 例如, 对于 x_1, 我们有

左端列	d_1 x_1	x_1 列 × 左端列
1	4	4×1
d_2	$-\frac{1}{4}$	$-\frac{1}{4}d_2$
d_3	$\frac{3}{2}$	$\frac{3}{2}d_3$
0	2	2×0
x_1 的简约费用 $=4-\frac{1}{4}d_2+\frac{3}{2}d_3-d_1$		

注意, 对基变量应用这些计算总是产生零简约费用, 这是一个已经证明的理论结果. 还有, 对于解列应用相同的规则, 得到 $z=1\,350+100d_2+230d_3$.

因为我们正在处理的是极大化问题, 所以只要对于所有的非基变量新简约费用 (z 方程的系数) 保持非负当前解就保持最优. 因此, 对于非基变量 x_1, x_4, x_5, 有下列**最优性条件**:

$$\begin{aligned} 4-\tfrac{1}{4}d_2+\tfrac{3}{2}d_3-d_1 &\geqslant 0\\ 1+\tfrac{1}{2}d_2 &\geqslant 0\\ 2-\tfrac{1}{4}d_2+\tfrac{1}{2}d_3 &\geqslant 0 \end{aligned}$$

这些条件必须同时满足才能维持当前最优解的最优性.

为解释这些条件的使用, 假定 TOYCO 的目标函数从

$$\max \quad z=3x_1+2x_2+5x_3$$

变为

$$\max \quad z=2x_1+x_2+6x_3$$

则 $d_1=2-3=-1$ 美元, $d_2=1-2=-1$ 美元, $d_3=6-5=1$ 美元. 代入到给定的条件中, 得到

$$4-\tfrac{1}{4}d_2+\tfrac{3}{2}d_3-d_1=4-\tfrac{1}{4}(-1)+\tfrac{3}{2}(1)-(-1)=6.75>0 \quad (\text{满足})$$

$$1+\tfrac{1}{2}d_2=1+\tfrac{1}{2}(-1)=0.5>0 \quad (\text{满足})$$

$$2-\tfrac{1}{4}d_2+\tfrac{1}{2}d_3=2-\tfrac{1}{4}(-1)+\tfrac{1}{2}(1)=2.75>0 \quad (\text{满足})$$

结果说明, 所作的上述改变仍保持当前解 ($x_1=0$, $x_2=100$, $x_3=230$) 最优. 因此, 除了目标函数变化到 $z=1\,350+100d_2+230d_3=1\,350+100\times(-1)+230\times 1=1\,480$ 美元, 不需要作进一步的计算. 如果有任何一个条件不满足的话, 则必须求出新的解 (见第 4 章).

到目前为止的讨论已处理了极大化的情况, 对于极小化情况, 差别仅是简约费用 (z 方程系数) 必须是 $\leqslant 0$ 才能维持最优性.

一般最优性条件可以用于确定特定的情况, 这里改变量 d_j 一次仅有一个发生改变, 而不是同步改变. 这种分析等价于考虑下列 3 种情况:

(1) $\max \ z = (3 + d_1)x_1 + 2x_2 + 5x_3$

(2) $\max \ z = 3x_1 + (2 + d_2)x_2 + 5x_3$

(3) $\max \ z = 3x_1 + 2x_2 + (5 + d_3)x_3$

单个条件可以看成是同步改变情况的特例[①].

情况 1 在同步改变的条件中置 $d_2 = d_3 = 0$, 即有

$$4 - d_1 \geqslant 0 \quad \Rightarrow -\infty < d_1 \leqslant 4$$

情况 2 在同步改变的条件中置 $d_1 = d_3 = 0$, 即有

$$\left.\begin{array}{ll} 4 - \frac{1}{4}d_2 \geqslant 0 & \Rightarrow d_2 \leqslant 16 \\ 1 + \frac{1}{2}d_2 \geqslant 0 & \Rightarrow d_2 \geqslant -2 \\ 2 - \frac{1}{4}d_2 \geqslant 0 & \Rightarrow d_2 \leqslant 8 \end{array}\right\} \Rightarrow -2 \leqslant d_2 \leqslant 8$$

情况 3 在同步改变的条件中置 $d_1 = d_2 = 0$, 即有

$$\left.\begin{array}{ll} 4 + \frac{3}{2}d_3 \geqslant 0 & \Rightarrow d_3 \geqslant -\frac{8}{3} \\ 2 + \frac{1}{2}d_3 \geqslant 0 & \Rightarrow d_3 \geqslant -4 \end{array}\right\} \Rightarrow -\frac{8}{3} \leqslant d_3 < \infty$$

所给出的单个条件能够依照总的单位收入来表示. 例如, 对于玩具卡车 (变量 x_2), 总的单位收入是 $2 + d_2$, 相应的条件是 $-2 \leqslant d_2 \leqslant 8$, 表示成

$$2 + (-2) \leqslant 2 + d_2 \leqslant 2 + 8$$

或

$$0 \text{ 美元} \leqslant \text{玩具卡车的单位收入} \leqslant 10 \text{ 美元}$$

这个条件假定玩具火车和玩具汽车的单位收入分别固定在 3 美元和 5 美元.

允许区间 [0 美元, 10 美元] 表明, 玩具卡车 (变量 x_2) 的单位收入可以是最低 0 美元或最高 10 美元而不改变当前的最优解, $x_1 = 0$, $x_2 = 100$, $x_3 = 230$. 但总收入将变化到 $1\,350 + 100d_2$.

注意到下列事实是重要的, 改变量 d_1, d_2, d_3 可以均在它们所许可的单个区域内, 而不必满足同步条件, 反之亦然. 例如, 考虑

$$\max \quad z = 6x_1 + 8x_2 + 3x_3$$

这里 $d_1 = 6 - 3 = 3$ 美元, $d_2 = 8 - 2 = 6$ 美元, $d_3 = 3 - 5 = -2$ 美元, 它们都在可允许的单个区域 ($-\infty < d_1 \leqslant 4$, $-2 \leqslant d_2 \leqslant 8$, $-\frac{8}{3} \leqslant d_3 < \infty$) 内. 然而, 相应的同步

① 单个区域是所有线性规划软件的标准输出. 同步条件并不是输出的内容, 大概是因为对于大问题它们过于麻烦.

改变条件将产生

$$4-\frac{1}{4}d_2+\frac{3}{2}d_3-d_1=4-\frac{1}{4}(6)+\frac{3}{2}(-2)-3=-3.5<0 \quad (\text{不满足})$$

$$1+\frac{1}{2}d_2=1+\frac{1}{2}(6)=4>0 \quad (\text{满足})$$

$$2-\frac{1}{4}d_2+\frac{1}{2}d_3=2-\frac{1}{4}(6)+\frac{1}{2}(-2)=-0.5<0 \quad (\text{不满足})$$

上述结果可以总结如下:

(1) 只要目标函数系数的改变量 $d_j, j=1,2,\cdots,n$ 满足所有的最优性条件 (当改变是同时发生时), 或落在最优性区域之内 (当改变是单个发生时), 那么变量的最优值保持不变.

(2) 对于其他的情况, 也就是同步改变时最优性条件不满足或者单个可行性区域被破坏, 那么可以用新的 d_j 值重解该问题, 或者应用将在第 4 章介绍的后最优分析来解决此类问题.

习题 3.6D①

1. 在 TOYCO 模型中, 确定在下列情况下当前解是否改变:

(a) $z=2x_1+x_2+4x_3$ **(b)** $z=3x_1+6x_2+x_3$

(c) $z=8x_1+3x_2+9x_3$

2. B&K 食品店出售 3 种类型的软饮料: 商标名称为 A1 可乐、A2 可乐和较廉价的店品牌 BK 可乐. 每听 A1 可乐、A2 可乐和 BK 可乐的价格分别是 80 美分、70 美分和 60 美分. 在通常情况下, 商店一天销售全部可乐不超过 500 听. 虽然 A1 可乐是公认的品牌, 但客户却愿意购买 A2 可乐和 BK 可乐, 因为它们便宜. 估计每天的销售至少是 100 听 A1 可乐, 并且 A2 可乐和 BK 可乐合在一起的销售量与 A1 销售量相比至少是 4:2.

(a) 证明最优解并不要求销售 BK 牌子的可乐.

(b) BK 牌子的可乐每听价格应该增加多少才值得被 B&K 销售?

(c) 为了与其他商店竞争, B&K 决定所有 3 种类型可乐的价格每听降价 5 美分. 重新计算简约费用, 确定这种促销是否会改变当前的最优解.

3. Baba 家具公司雇用 4 位木匠在 10 天中装配桌子和椅子. 花费 2 人小时装配一张桌子, 0.5 人小时装配一把椅子. 顾客通常买 1 张桌子和 4~6 把椅子. 每张桌子的价格是 135 美元, 每把椅子的价格是 50 美元. 公司实行每天 8 小时工作制.

(a) 确定 10 天的最优生产策略.

(b) 如果当前每张桌子和每把椅子的单位价格减少 10%, 用灵敏度分析确定由 (a) 获得的最优解是否发生改变.

(c) 如果当前每张桌子和每把椅子的价格分别变到 120 美元和 25 美元, 在 (a) 中的最优解会发生改变吗?

4. Elkins 银行计划在下个月分配最大额度为 200 000 美元的个人和汽车贷款. 银行对个人贷款的利息是 14%, 对汽车贷款的利息是 12%. 两种类型的贷款要在一年后偿还. 经验表明,

---

① 在这组习题中, 你可以发现, 借助于 TORA, 能够非常方便地生成最优单纯形表.

大约有 3% 的个人贷款和 2% 的汽车贷款不能归还. 银行通常分配给汽车贷款金额至少是个人贷款的 2 倍.

(a) 确定两种贷款之间的最优基金分配和所有返还贷款的净利润.

(b) 如果个人贷款和汽车贷款的利率分别变化 4% 和 3%, 用灵敏度分析确定在 (a) 中的最优解是否将会发生变化.

***5.** Electra 公司生产 4 种类型的马达, 每种马达在各自的装配线上装配. 装配线各自的生产能力是每天 500, 500, 800, 750 台马达. 1 型马达使用 8 个单位的某种电子元件, 2 型马达使用 5 个单位, 3 型马达使用 4 个单位, 4 型马达使用 6 个单位. 电子元件的供应者每天能够提供 8 000 件. 对于每种类型的马达, 每台的价格分别是 60 美元、40 美元、25 美元和 30 美元.

(a) 确定每天的最优产品生产策略.

(b) 现有的生产计划安排能满足 Electra 的需要. 然而, 由于竞争, Electra 可能需要降低 2 型马达的价格. 要想不改变现有的生产计划安排最多可降价多少?

(c) Electra 还决定大幅度把所有类型马达的价格削减 25%. 用灵敏度分析确定是否最优解保持不变.

(d) 目前, 4 型马达不生产. 它的价格应该在生产计划中增加多少?

6. Popeye 罐头厂已签约以每磅 7 美分的价格每天收购 60 000 磅成熟的西红柿, 用它生产罐头番茄汁、番茄酱和番茄膏. 将罐头产品包装成 24 听一箱. 一听番茄汁使用 1 磅的新鲜西红柿, 一听番茄酱使用 $\frac{1}{2}$ 磅, 一听番茄膏使用 $\frac{3}{4}$ 磅. 公司的每天的市场份额限制在 2 000 箱番茄汁、5 000 箱番茄酱和 6 000 箱番茄膏. 每箱番茄汁、番茄酱和番茄膏的批发价格分别是 21 美元、9 美元和 12 美元.

(a) 为 Popeye 罐头厂制定一个最优的日生产规划.

(b) 如果每箱番茄汁和番茄膏的价格作为问题的已知值保持固定不变, 用灵敏度分析确定, 在保持最优产品生产规划不变的情况下, Popeye 公司对一箱番茄酱收费的单位价格区间.

7. Dean 家具公司用预制木材装配普通和豪华厨房餐具柜. 普通橱柜被涂成白色, 豪华橱柜被上漆粉饰. 着色和上漆在一个车间完成. 装配车间每天的生产能力是 200 个普通橱柜和 150 个豪华橱柜. 豪华橱柜上漆花费的单位时间是普通橱柜着色的 2 倍. 如果着色/上漆车间只给豪华橱柜上漆, 它每天能够完成 180 套. 公司估计每套普通橱柜和豪华橱柜的收入分别是 100 美元和 140 美元.

(a) 为该问题建立一个线性规划, 并求每天的最优生产计划安排.

(b) 假定竞争迫使每套普通橱柜和每套豪华橱柜的价格减少到 80 美元. 用灵敏度分析确定 (a) 中的最优解是否保持不变.

8. 100% *最优性规则*. 规则类似于在习题 3.6C 中第 12 题概述的100% *可行性规则*, 它还用来检验在当前解的最优性条件下同时把所有 c_j 改变到 c_j+d_j, $j=1,2,\cdots,n$ 的效果. 假定 $u_j \leqslant d_j \leqslant v_j$ 是使用 3.6.3 节的方法, 一次改变一个值由 c_j 变化到 c_j+d_j 而获得的最优性区域. 在这种情况下, $u_j \leqslant 0$ $(v_j \geqslant 0)$, 因为它表示保持当前解最优的 c_j 的最大允许的减少 (增加) 值. 对于 $u_j \leqslant d_j \leqslant v_j$ 的情况, 定义 r_j 等于 $\frac{d_j}{v_j}$(如果 d_j 是正的), 或等于 $\frac{d_j}{u_j}$(如果 d_j 是负的). 由定义, $0 \leqslant r_j \leqslant 1$. 100% 规则是说, 对于当前解保持最优的充

分 (但不必要) 条件是 $r_1 + r_2 + \cdots + r_n \leqslant 1$. 如果条件不满足, 当前解可能保持或者可能不保持最优. 如果 d_j 落在指定的区域之外, 则规则不能使用.

以下列两种情况为例, 论述 100% 最优性规则很弱, 不足以作为可靠的决策工具:

(a) 第 1 题的第 (b) 部分和第 (c) 部分.

(b) 第 7 题的 (b) 部分.

3.6.4 用 TORA、Excel 规划求解和 AMPL 作灵敏度分析

我们现在有了需要解读由线性规划软件提供的输出的所有工具, 特别是关于灵敏度分析的工具. 我们将用 TOYCO 例子来演示 TORA、Excel 规划求解和 AMPL 的输出.

TORA 的线性规划输出报告自动提供了灵敏度分析的数据, 如图 3.14 所示 (文件 toraTOYCO.txt). 该输出包含了简约费用和对偶价格, 同时提供了所允许的最优性区域与可行性区域.

Sensitivity Analysis

Variable	CurrObjCoeff	MinObjCoeff	MaxObjCoeff	Reduced Cost
x1:	3.00	-infinity	7.00	4.00
x2:	2.00	0.00	10.00	0.00
x3:	5.00	2.33	infinity	0.00
Constraint	**Curr RHS**	**Min RHS**	**Max RHS**	**Dual Price**
1(<):	430.00	230.00	440.00	1.00
2(<):	460.00	440.00	860.00	2.00
3(<):	420.00	400.00	infinity	0.00

图 3.14 TOYCO 模型的 TORA 灵敏度分析

图 3.15 提供了 Excel规划求解的 TOYCO 模型 (文件 TOYCO.xls) 和它的灵敏度分析报告. 单击“**规划求解参数**”对话框中“ 求解 (S) ”后, 新的对话框“**规划求解结果**”将让你有机会获得关于解的更详细资料, 包含重要的灵敏度分析报告[①]. 该报告存在于另一张 Excel 电子表格中, 可以在屏幕的底部选择显示它. 然后可以单击**敏感性报告 1**来查看灵敏度分析的结果. 这个报告类似于 TORA 的报告, 但有 3 处不同: (1) 简约费用的符号相反; (2) 用名称*阴影价格*[②] 代替名称*对偶价格*; (3) 最优性区域是针对改变量 d_j 和 D_i 的, 而不是针对总目标系数和约束右端项的. 这些差别并不重要, 关键是对结果的解释是一样的.

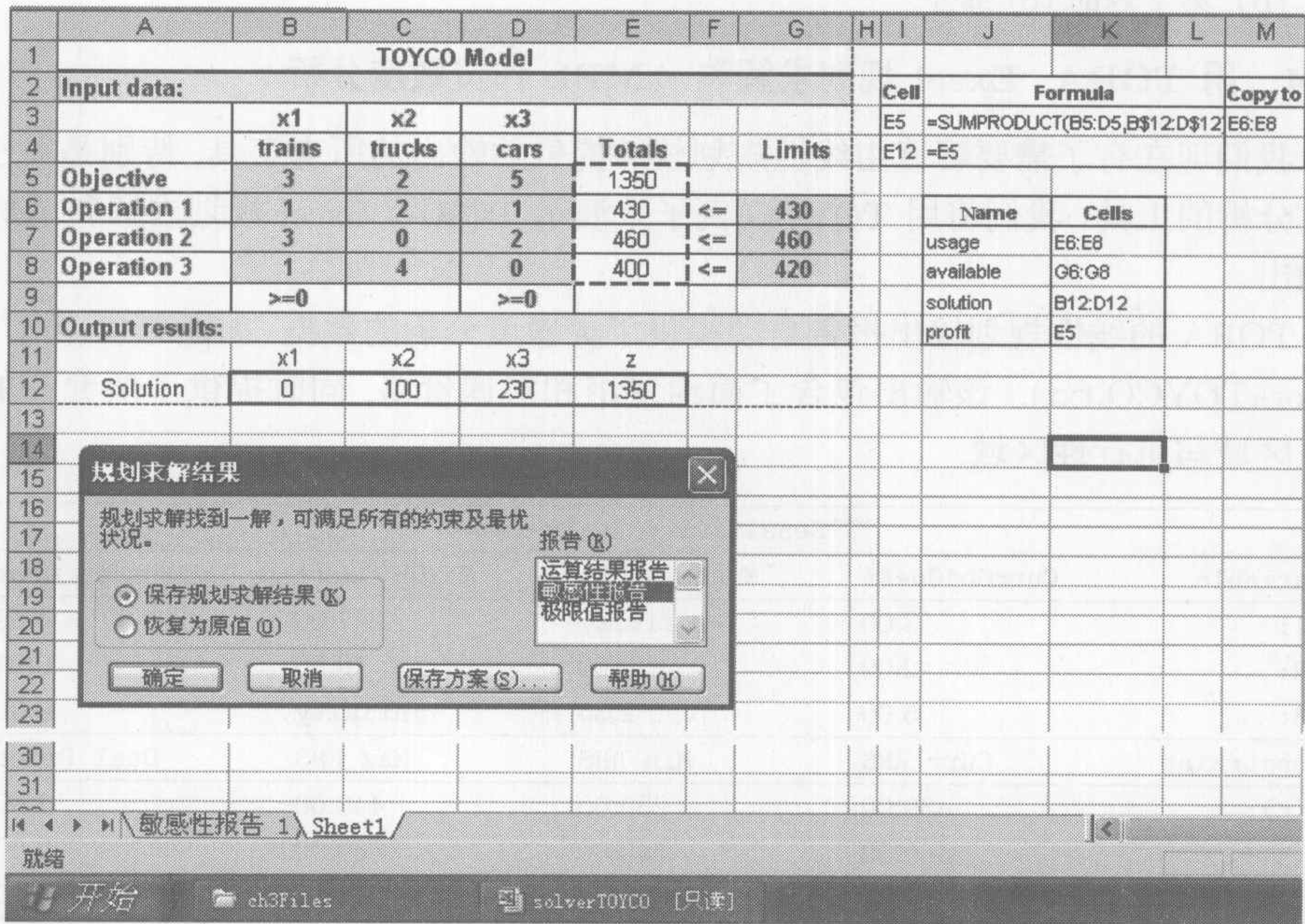
在 AMPL 中, 也可以提供灵敏度分析报告. 文件 amplTOYCO.txt 提供了确定灵敏度分析报告所需的编码. 它还需要下列语句:

```
option solver cplex;
option cplex_options 'sensitivity';
```

① 在 Excel 规划求解结果对话框中显示的是“敏感性报告”. —— 译者注

② 这是中文 Excel 表显示的名称, 通常译为“影子价格”. —— 译者注

```
solve;
#-------------------------------sensitivity analysis
display oper.down,oper.current,oper.up,oper.dual>a.out;
display x.down,x.current,x.up,x.rc>a.out;
```



	A	B	C	D	E	F	G
1	TOYCO Model						
2	Input data:						
3		x1	x2	x3			
4		trains	trucks	cars	Totals		Limits
5	Objective	3	2	5	1350		
6	Operation 1	1	2	1	430	<=	430
7	Operation 2	3	0	2	460	<=	460
8	Operation 3	1	4	0	400	<=	420
9		>=0	>=0	>=0			
10	Output results:						
11		x1	x2	x3	z		
12	Solution	0	100	230	1350		

Cell	Formula	Copy to
E5	=SUMPRODUCT(B5:D5,B$12:D$12)	E6:E8
E12	=E5	

Name	Cells
usage	E6:E8
available	G6:G8
solution	B12:D12
profit	E5

可变单元格

单元格	名字	终值	递减成本	目标式系数	允许的增量	允许的减量
$B$12	Solution x1	0	-4	3	4	1E+30
$C$12	Solution x2	100	0	2	8	2
$D$12	Solution x3	230	0	5	1E+30	2.666666667

约束

单元格	名字	终值	阴影价格	约束限制值	允许的增量	允许的减量
$E$6	Operation 1 Totals	430	1	430	10	200
$E$7	Operation 2 Totals	460	2	460	400	20
$E$8	Operation 3 Totals	400	0	420	1E+30	20

图 3.15　TOYCO 模型的 Excel 规划求解敏感性报告

需要 CPLEX option 语句才能够获得标准的灵敏度分析报告. 在 TOYCO 模型中, 指标变量和约束使用的根名称分别为 x 和 oper, 用这些名称, display 语句中的提示性后缀 .down、.current 和 .up 自动产生如图 3.16 所示的编排好的灵敏度分析报告. 后缀 .dual 和 .rc 提供了对偶价格和简约费用.

```
:  oper.down    oper.current        oper.up    oper.dual   :=
1     230          430                440          1
2     440          460                860          2
3     400          420               1e+20         0
;

:        x.down      x.current     x.up        x.rc       :=
1    -1e+20          3             7           -4
2        0           2            10            0
3     2.33333        5           1e+20          0
;
```

图 3.16 TOYCO 模型的 AMPL 灵敏度分析报告

AMPL 软件的另一种标准灵敏度分析报告实际上是要求解给定目标系数和约束右端项的值均在一定区域的线性规划模型. 通过 commands 命令的使用, AMPL 自动运行这个过程 (见附录 A.7). 以 TOYCO 模型即文件 amplTOYCO.txt 为先决条件, 我们打算研究改变 b[1](也就是操作 1 总可用时间) 后的效果. 为此可以通过把 solve 和 display 从 amplTOYCO.txt 移动到一个新文件 (不妨命名为analysis.txt):

```
repeat while b[1]<=500
{
solve;
display z, x;
let b[1]:=b[1]+1;
};
```

接下来, 在 ampl 提示下输入下面的语句:

```
ampl:  model amplTOYCO.txt;
ampl:  commands analysis.txt;
```

第 1 行提供模型和它的数据, 第 2 行提供 b[1] 以 430 为初始值 (初始值由 amplTOYCO.txt 给出) 并连续加 1 直到 b[1] 达到 500 为止的最优解. 分析这个输出结果可以让我们对最优解关于 b[1] 的变化的灵敏度进行分析. 对于其他的系数 (包含同步变化的情况), 也可以采用类似的方法.

习题 3.6E ①

1. 考虑习题 2.3C(第 2 章) 的第 1 题. 用对偶价格决定是否值得增加第 4 年的基金.
2. 考虑习题 2.3C(第 2 章) 的第 2 题.

① 在回答这组习题之前, 最好先用 APML、Excel 规划求解和 TORA 生成灵敏度分析报告.

(a) 用对偶价格确定投资的全部收益.

(b) 如果你希望第 1 年年末在娱乐方面花掉 1 000 美元, 它将如何影响第 5 年年初的累积基金?

3. 考虑习题 2.3C(第 2 章) 的第 3 题.

(a) 给出模型对偶价格的经济学解释.

(b) 说明：相应于第三季度初借钱上界的对偶价格, 如何可以由一年中五个指定日期的进出现金流平衡方程所对应的对偶价格得到.

4. 考虑习题 2.3C(第 2 章) 的第 4 题. 用对偶价格确定每年的收益率.

***5.** 考虑习题 2.3C(第 2 章) 的第 5 题. 用对偶价格确定商业主管在这项计划中投资更多的资金是否值得.

6. 考虑习题 2.3C(第 2 章) 的第 6 题. 用对偶价格确定赌徒赌更多的钱是否可取.

7. 考虑习题 2.3D(第 2 章) 的第 1 题. 叙述模型单位产品费用的对偶价格.

8. 考虑习题 2.3D(第 2 章) 的第 2 题. 假定机器 1 和机器 2 任何能力增长只能通过加班来获得. 对于两台机器的任一台, 公司能够承受的每小时最大费用是多少?

***9.** 考虑习题 2.3D(第 2 章) 的第 3 题.

(a) 假定制造商能够以每单位 12 美元购买更多的原料 A. 这样做可取吗?

(b) 你愿意推荐制造商以每单位 5 美元价格购买更多的原料 B 吗?

10. 考虑习题 2.3E(第 2 章) 的第 10 题.

(a) 哪一个规格约束对最优解有不利的影响?

(b) 公司为每种矿石每吨最多应该付多少钱?

参 考 文 献

Bazaraa, M., J. Jarvis, and H. Sherali, *Linear Programming and Network Flows*, 2nd ed.,Wiley,New York, 1990.

Dantzig, G., *Linear Programming and Extensions*, Princeton University Press, Princeton,NJ, 1963.

Dantzig, G., and M. Thapa, *Linear Programming 1: Introduction*, Springer, New York, 1997.

Fourer, R.,D. Gay, and B. Kernighan, *AMPL, A Modeling Language for Mathematical Programming*, 2nd ed., Brooks/Cole-Thompson, Pacific Grove, CA, 2003.

Nering, E., and A. Tucker, *Linear Programming and Related Problems*, Academic Press, Boston, 1992.

Schrage, L., *Optimization Modeling with LINGO*, LINDO Systems, Inc., Chicago, 1999.

Taha, H., "Linear Programming," Chapter II-1 in *Handbook of Operations Research*, J. Moder and S. Elmaghraby (eds.), Van Nostrand Reinhold, New York, 1987.

第 4 章　对偶性与后最优分析

本章导读　第 3 章通过确定使最优基本解不变的模型参数的范围, 讲解了最优解的灵敏度分析. 接下来, 一个很自然的问题就是*后最优分析*, 其目的是通过让模型参数有目的的改变来求出新的最优解. 尽管后最优分析能够用 3.6 节单纯形计算来完成, 但这一章完全基于对偶问题.

作为最低的要求, 你需要学习对偶问题及其经济学解释 (4.1 节、4.2 节和 4.3 节). 在 4.1 节, 对偶问题的数学定义非常抽象. 然而, 学到 4.3 节时, 你将看到对偶问题引导出的线性规划模型有着非常有意思的经济学解释, 包括*对偶价格*和*简约费用*. 这还为新的*对偶单纯形算法*(它是后最优分析的先决条件) 的发展奠定基础. 还可以将对偶单纯形算法用于第 8 章的整数规划.

4.4.2 节中的*广义单纯形算法*意在表明单纯形法并不死板, 这意味着, 你可以修改某些规则来处理开始时既不可行又非最优的一些问题. 然而, 可以将这段内容跳过, 而不失阅读全书的连续性.

可以用 TORA 的交互模式来加强对对偶单纯形法计算细节的理解.

本章包括 14 个例子、56 个节后习题和 2 个案例. 案例在本书 (下册) 的附录 E 中.

4.1　对偶问题的定义

对偶 (dual) 问题是由**原始** (primal) 线性规划模型直接按系统化定义的一种线性规划. 这两个问题有着如此紧密的关系, 以至于一个问题的最优解自动地提供另一个问题的最优解.

在大多数线性规划的处理中, 对偶问题是由各种形式的原始问题来定义的, 这些形式包括最优化含义 (极大化或极小化)、约束的类型 ($\leqslant$、$\geqslant$ 或 $=$), 以及变量的方向 (非负、无限制). 这种处理有点混乱, 针对这个原因, 我们给出一种*单一*的定义, 把所有形式的原始问题自动地包含进来.

我们关于对偶问题的定义要求将原始问题表示成 3.1 节提出的*等式约束形式* (所有约束是等式方程, 有非负的右端项, 并且所有变量非负). 这种要求与初始单纯形表的形式是一致的. 因此, 从原始问题最优解得到的任何结果都可以直接用到相应的对偶问题上.

为了说明如何构造对偶问题, 定义等式形式的原始问题如下:

$$\max\ (\text{或}\ \min) \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1, 2, \cdots, m$$

$$x_j \geqslant 0, \quad j = 1, 2, \cdots, n$$

变量 $x_j, j = 1, 2, \cdots, n$ 包含任何剩余变量、松弛变量和人工变量.

表 4.1 说明如何从原始问题构造出对偶问题. 事实上, 我们有

(1) 对偶变量是针对原始问题每个 (约束) 方程定义的.

(2) 对偶约束是针对原始问题的每个变量定义的.

(3) 原始问题变量约束 (列) 的系数定义对偶约束左端项的系数, 它的目标系数定义了右端项.

(4) 对偶目标系数等于原始问题约束方程的右端项.

表 4.1　从原始问题构造对偶问题

对偶变量	原始问题变量						右端项
	x_1	x_2	$\cdots$	x_j	$\cdots$	x_n	
	c_1	c_2	$\cdots$	c_j	$\cdots$	c_n	
y_1	a_{11}	a_{12}	$\cdots$	a_{1j}	$\cdots$	a_{1n}	b_1
y_2	a_{21}	a_{22}	$\cdots$	a_{2j}	$\cdots$	a_{2n}	b_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	a_{m1}	a_{m2}	$\cdots$	a_{mj}	$\cdots$	a_{mn}	b_m
				↑ 第 j 个 对偶约束			↑ 对偶目标 函数系数

表 4.2 概括了确定最优化含义 (极大化或极小化)、约束的类型 ($\leqslant$、$\geqslant$ 或 $=$), 以及对偶变量符号的规则. 注意, 在对偶问题中最优化含义总是与原始问题最优化含义相反. 记住对偶问题约束类型 (也就是, $\leqslant$ 或 $\geqslant$) 的一种容易的方法是, 如果对偶目标是求极小, 则所有约束都是 $\geqslant$ 型的. 当对偶目标是求极大时, 约束类型相反.

表 4.2　构造对偶问题的规则

原始问题的目标 *	对偶问题		
	目标	约束类型	变量符号
极大化	极小化	$\geqslant$	无限制
极小化	极大化	$\leqslant$	无限制

* 所有原始问题约束是等式方程, 具有非负右端项, 并且所有变量非负.

下面的例子说明表 4.2 规则的使用, 并说明我们的定义自动适用于所有形式的原始问题.

例 4.1-1

原始问题	等式方程形式的原始问题	对偶变量
max $z=5x_1+12x_2+4x_3$	max $z=5x_1+12x_2+4x_3+0x_4$	
s.t. $x_1+2x_2+x_3\leqslant 10$	s.t. $x_1+2x_2+x_3+x_4=10$	y_1
$2x_1-x_2+3x_3=8$	$2x_1-x_2+3x_3+0x_4=8$	y_2
$x_1,x_2,x_3\geqslant 0$	$x_1,x_2,x_3,x_4\geqslant 0$	

对偶问题

$$
\begin{aligned}
\min\quad & w=10y_1+8y_2\\
\text{s.t.}\quad & y_1+2y_2\geqslant 5\\
& 2y_1-y_2\geqslant 12\\
& y_1+3y_2\geqslant 4\\
& \left.\begin{array}{l} y_1+0y_2\geqslant 0\\ y_1,y_2\text{无限制}\end{array}\right\}\Rightarrow(y_1\geqslant 0,y_2\text{无限制})
\end{aligned}
$$

例 4.1-2

原始问题	等式方程形式的原始问题	对偶变量
min $z=15x_1+12x_2$	min $z=15x_1+12x_2+0x_3+0x_4$	
s.t. $x_1+2x_2\geqslant 3$	s.t. $x_1+2x_2-x_3+0x_4=3$	y_1
$2x_1-4x_2\leqslant 5$	$2x_1-4x_2+0x_3+x_4=5$	y_2
$x_1,x_2\geqslant 0$	$x_1,x_2,x_3,x_4\geqslant 0$	

对偶问题

$$
\begin{aligned}
\max\quad & w=3y_1+5y_2\\
\text{s.t.}\quad & y_1+2y_2\leqslant 15\\
& 2y_1-4y_2\leqslant 12\\
& \left.\begin{array}{l} -y_1\leqslant 0\\ y_2\leqslant 0\\ y_1,y_2\text{无限制}\end{array}\right\}\Rightarrow(y_1\geqslant 0,y_2\leqslant 0)
\end{aligned}
$$

例 4.1-3

原始问题	等式方程形式的原始问题	对偶变量
	替换 $x_1=x_1^+-x_1^-$	
max $z=5x_1+6x_2$	max $z=5x_1^+-5x_1^-+6x_2$	
s.t. $x_1+2x_2=5$	s.t. $x_1^- - x_1^+ + 2x_2=5$	y_1
$-x_1+5x_2\geqslant 3$	$-x_1^- + x_1^+ + 5x_2 - x_3=3$	y_2
$4x_1+7x_2\leqslant 8$	$4x_1^- - 4x_1^+ + 7x_2 + x_4=8$	y_3
x_1无限制, $x_2\geqslant 0$	$x_1^-,x_1^+,x_2,x_3,x_4\geqslant 0$	

对偶问题

$$
\begin{aligned}
\min\quad & z = 5y_1 + 3y_2 + 8y_3 \\
\text{s.t.}\quad & \left.\begin{aligned} y_1 - y_2 + 4y_3 &\geqslant 5 \\ -y_1 + y_2 - 4y_3 &\geqslant -5 \end{aligned}\right\} \Rightarrow (y_1 - y_2 + 4y_3 = 5) \\
& 2y_1 + 5y_2 + 7y_3 \geqslant 6 \\
& \left.\begin{aligned} -y_2 &\geqslant 0 \\ y_3 &\geqslant 0 \\ y_1, y_2, y_3 & \text{无限制} \end{aligned}\right\} \Rightarrow (y_1\text{无限制}, y_2 \leqslant 0, y_3 \geqslant 0)
\end{aligned}
$$

第 1 个和第 2 个约束用等式方程来代替. 这种情况发生的一般规则是, 无限制原始问题变量总是对应于对偶问题的等式约束. 相反地, 原始问题的等式约束总对应于无限制对偶变量, 正如第 1 个原始问题约束显示的那样.

构造对偶问题规则的总结 从前面的例子中得到的一般结论是, 原始问题和对偶问题中的变量和约束是由表 4.3 所示的规则定义的. 我们可以检验, 这些明确的对应关系包含在表 4.2 所示的一般规则中.

表 4.3 构造对偶问题的规则

极大化问题		极小化问题
约束		变量
$\geqslant$	$\Leftrightarrow$	$\leqslant 0$
$\leqslant$	$\Leftrightarrow$	$\geqslant 0$
$=$	$\Leftrightarrow$	无限制
变量		约束
$\geqslant 0$	$\Leftrightarrow$	$\geqslant$
$\leqslant 0$	$\Leftrightarrow$	$\leqslant$
无限制	$\Leftrightarrow$	$=$

注意, 此表并不用于指定原始问题和对偶问题. 这里重要的是最优化含义. 如果原始问题是极大化, 则对偶问题就是极小化, 反之亦然.

习题 4.1A

1. 在例 4.1-1 中, 假如把原始问题最优化含义改为求极小, 导出相应的对偶问题.

***2.** 在例 4.1-2 中, 在原始问题中增加第 3 个约束 $3x_1 + x_2 = 4$, 导出相应的对偶问题.

3. 在例 4.1-3 中, 证明即使原始问题最优化含义改为求极小, 无限制原始问题变量总是对应于等式对偶约束.

4. 写出下列每个原始问题的对偶问题.

(a) max $z = -5x_1 + 2x_2$

$$\begin{aligned} \text{s.t.}\quad & -x_1 + x_2 \leqslant -2 \\ & 2x_1 + 3x_2 \leqslant 5 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

(b) min $z = 6x_1 + 3x_2$

$$\begin{aligned} \text{s.t.}\quad & 6x_1 - 3x_2 + x_3 \geqslant 2 \\ & 3x_1 + 4x_2 + x_3 \geqslant 5 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

*(c) max $z = x_1 + x_2$

$$\begin{aligned} \text{s.t.}\quad & 2x_1 + x_2 = 5 \\ & 3x_1 - x_2 = 6 \\ & x_1, x_2 \text{无限制} \end{aligned}$$

*5. 考虑例 4.1-1. 单纯形法应用于原始问题要求在标准原始问题的第 2 个约束中使用人工变量, 以保证初始基本解. 证明: 原始问题等式形式中人工变量的出现并不影响对偶的定义, 因为它导出多余的对偶约束.

6. 判断正误:

(a) 对偶问题的对偶得到原始问题.

(b) 如果原始问题约束最初是等式形式, 相应的对偶变量必定无限制.

(c) 如果原始问题约束的类型是 $\leqslant$, 且原始问题目标是求极大 (极小), 则相应的对偶变量一定非负 (非正).

(d) 如果原始问题约束的类型是 $\geqslant$, 且原始问题目标是求极小 (极大), 则相应的对偶变量一定非负 (非正).

(e) 一个无限制的原始问题变量将导致一个等式对偶约束.

4.2 原始–对偶关系

原始线性规划模型中的变化将改变当前最优表的元素, 这反过来又会影响到当前解的最优性和 (或) 当前解的可行性. 本节介绍一些原始–对偶之间的关系, 用来重新计算最优单纯形表的元素. 这些关系将构成线性规划模型经济学解释和后最优性分析的基础.

本节先简要地复习矩阵, 因为它是一种完成单纯形表计算的便利工具.

4.2.1 简单矩阵运算的复习

单纯形表的计算只用到 3 种初等矩阵运算: (行向量)×(矩阵), (矩阵)×(列向量), 以及 (标量)×(矩阵). 在这里概述这些运算是为了方便. 首先, 我们介绍一些矩阵的定义 ①:

(1) 一个维数为 $(m \times n)$ 的**矩阵$\boldsymbol{A}$** 是一个具有 m 行 n 列元素的矩形阵列.

(2) 一个维数为 m 的**行向量$\boldsymbol{V}$** 是一个 $(1 \times m)$ 维的矩阵.

(3) 一个维数为 n 的**列向量$\boldsymbol{P}$** 是一个 $(n \times 1)$ 维的矩阵.

① 本书 (下册) 的附录 D 提供有更全面的矩阵复习.

这些定义的数学表达式如下：

$$\boldsymbol{V}=(v_1,v_2,\cdots,v_m),\ \boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},\ \boldsymbol{P}=\begin{pmatrix}p_1\\p_2\\\vdots\\p_n\end{pmatrix}$$

(1) (行向量 × 矩阵, $\boldsymbol{VA}$). 只有当行向量的维数等于 $\boldsymbol{A}$ 的行数, 该运算才有定义. 在这种情况下,

$$\boldsymbol{VA}=\left(\sum_{i=1}^m v_i a_{i1},\sum_{i=1}^m v_i a_{i2},\cdots,\sum_{i=1}^m v_i a_{in}\right)$$

例如,

$$\begin{aligned}(11,22,33)\begin{pmatrix}1&2\\3&4\\5&6\end{pmatrix}&=(1\times11+3\times22+5\times33,2\times11+4\times22+6\times33)\\&=(242,308)\end{aligned}$$

(2) (矩阵 × 列向量, $\boldsymbol{AP}$). 只有当 $\boldsymbol{A}$ 的列数等于列向量 $\boldsymbol{P}$ 的维数, 运算才有定义. 在这种情况下,

$$\boldsymbol{AP}=\begin{pmatrix}\sum_{j=1}^n a_{1j}p_j\\\sum_{j=1}^n a_{2j}p_j\\\vdots\\\sum_{j=1}^n a_{mj}p_j\end{pmatrix}$$

例如, 我们有

$$\begin{pmatrix}1&3&5\\2&4&6\end{pmatrix}\begin{pmatrix}11\\22\\33\end{pmatrix}=\begin{pmatrix}1\times11+3\times22+5\times33\\2\times11+4\times22+6\times33\end{pmatrix}=\begin{pmatrix}242\\308\end{pmatrix}$$

(3) (标量 × 矩阵, $\alpha\boldsymbol{A}$). 已知标量 (或常数) α, 乘法运算 $\alpha\boldsymbol{A}$ 将得到与 $\boldsymbol{A}$ 相同维数的矩阵, 其中第 (i,j) 元素等于 αa_{ij}. 例如, 已知 $\alpha=10$, 则有

$$(10)\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}=\begin{pmatrix}10&20&30\\40&50&60\end{pmatrix}$$

一般而言, $\alpha\boldsymbol{A}=\boldsymbol{A}\alpha$. 相同的运算等价地扩充到标量与向量的乘法. 例如, $\alpha\boldsymbol{V}=\boldsymbol{V}\alpha$ 和 $\alpha\boldsymbol{P}=\boldsymbol{P}\alpha$.

习题 4.2A

1. 考虑下面的矩阵:

$$A=\begin{pmatrix}1&2\\3&4\\5&6\end{pmatrix},\ P_1=\begin{pmatrix}1\\2\end{pmatrix},\ P_2=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

$$V_1=(11,22),\quad V_2=(-1,-2,-3)$$

在下面的每种情况中, 指出所给出的矩阵运算是否合法. 如果是, 请计算出结果.

*(a) AV_1 (b) AP_1 (c) AP_2 (d) V_1A *(e) V_2A (f) P_1P_2 (g) V_1P_1

4.2.2 单纯形表的布局图

第 3 章按照一种特殊形式设置了单纯形表. 这种特殊形式是本章内容的基础.

图 4.1 给出了*初始*和*一般*单纯形表的一种表达形式. 初始表中, 在初始变量下方的约束系数构成一个**单位矩阵** (identity matrix; 所有主对角元素等于 1, 且非对角元素为 0). 由于这种排列, 由高斯–若尔当行运算 (见第 3 章) 产生的后继单纯形表迭代将会修改这个单位矩阵的元素, 得到所谓的**逆矩阵** (inverse matrix). 在本章的余下部分我们将会看到, 逆矩阵是计算相应单纯形表所有元素的关键.

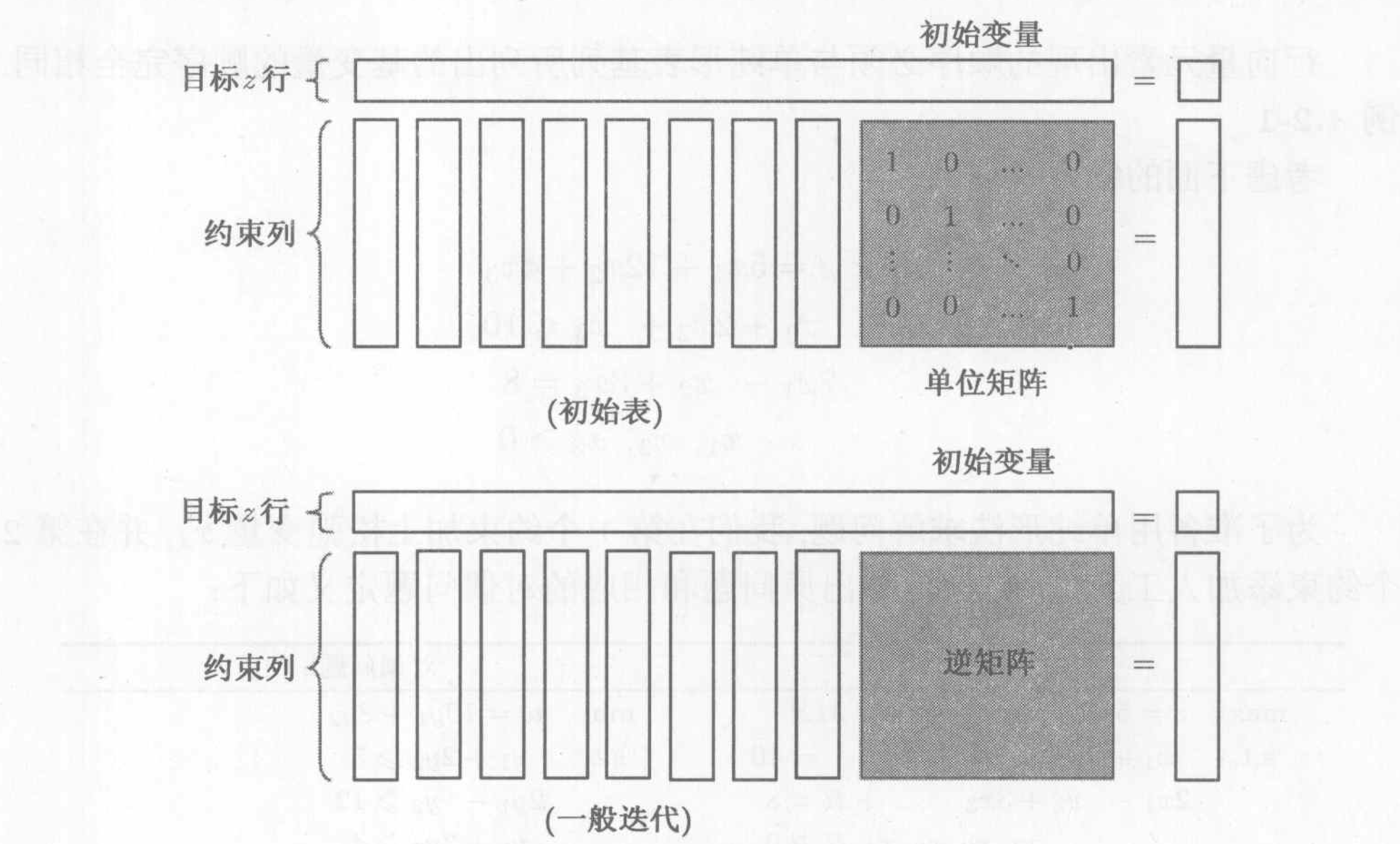


图 4.1 初始和一般单纯形表的示意表达

习题 4.2B

1. 考虑例 3.3-1 的最优单纯形表.

*(a) 找出最优逆矩阵.

(b) 证明右端项等于逆矩阵乘上原始约束的原始右端向量.

2. 对于例 3.4-1 的最后一张表解决第 1 题.

4.2.3　最优对偶解

原始问题与对偶问题的解有着如此紧密的关系, 以至于任何一个问题的最优解直接产生 (几乎不用再计算) 另一个问题的最优解. 因此, 在线性规划模型中, 变量的个数比约束的个数小得多, 通过求解对偶问题可以节省计算量, 因为从对偶问题的解可以自动地求出原始问题的解. 这个结果成立是因为单纯形的计算量在很大程度上 (尽管不是全部) 依赖于约束的个数 (见习题 4.2C 的第 2 题).

本节提供确定对偶变量的两种方法. 注意, 对偶问题的对偶是原始问题本身, 这意味着对偶问题的解还可以用来自动地生成原始问题的最优解.

方法 1

对偶变量 y_i 的最优值 = 初始变量 x_i 的最优原始 z 系数 + x_i 的原始目标系数

方法 2

对偶变量的最优值 = 最优原始基变量的原目标系数的行向量 × 最优原始逆矩阵

行向量元素出现的顺序必须与单纯形表基列所列出的基变量的顺序完全相同.

例 4.2-1

考虑下面的线性规划:

$$\begin{aligned}
\max\ & z = 5x_1 + 12x_2 + 4x_3 \\
\text{s.t.}\ & x_1 + 2x_2 + x_3 \leqslant 10 \\
& 2x_1 - x_2 + 3x_3 = 8 \\
& x_1,\ x_2,\ x_3 \geqslant 0
\end{aligned}$$

为了准备用单纯形法求解问题, 我们在第 1 个约束加上松弛变量 x_4, 并在第 2 个约束添加人工变量 R. 因此导出原问题和相应的对偶问题定义如下:

原始问题		对偶问题	
max	$z = 5x_1 + 12x_2 + 4x_3 - MR$	min	$w = 10y_1 + 8y_2$
s.t.	$x_1 + 2x_2 + x_3 + x_4 = 10$	s.t.	$y_1 + 2y_2 \geqslant 5$
	$2x_1 - x_2 + 3x_3 + R = 8$		$2y_1 - y_2 \geqslant 12$
	$x_1, x_2, x_3, x_4, R \geqslant 0$		$y_1 + 3y_2 \geqslant 4$
			$y_1 \geqslant 0$
			$y_2 \geqslant -M (\Rightarrow y_2$ 无限制$)$

表 4.4 提供了原始问题的最优单纯形表.

表 4.4 例 4.2-1 原始问题的最优单纯形表

基	x_1	x_2	x_3	x_4	R	解
z	0	0	$\frac{3}{5}$	$\frac{29}{5}$	$-\frac{2}{5}+M$	$54\frac{4}{5}$
x_2	0	1	$-\frac{1}{5}$	$\frac{2}{5}$	$-\frac{1}{5}$	$\frac{12}{5}$
x_1	1	0	$\frac{7}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{26}{5}$

现在说明如何用本节开始描述的两种方法来确定最优的对偶值.

方法 1 在表 4.4 中, 初始原始变量 x_4 和 R 分别唯一地对应于对偶变量 y_1 和 y_2. 因此, 我们确定最优对偶解如下:

初始原始基变量	x_4	R
z 方程的系数	$\frac{29}{5}$	$-\frac{2}{5}+M$
原始目标系数	0	$-M$
对偶变量	y_1	y_2
最优对偶值	$\frac{29}{5}+0=\frac{29}{5}$	$-\frac{2}{5}+M+(-M)=-\frac{2}{5}$

方法 2 最优逆矩阵, 即表 4.4 中初始变量 x_4 和 R 下面突出显示的部分, 即

$$\text{最优逆矩阵}=\begin{pmatrix}\frac{2}{5} & -\frac{1}{5}\\ \frac{1}{5} & \frac{2}{5}\end{pmatrix}$$

首先, 我们注意到, 最优原始变量列在表中, 其行的顺序为先 x_2 再 x_1. 这意味着两个变量原始目标系数的元素必须以相同的顺序出现, 也就是,

$$\begin{aligned}(\text{原始目标系数})&=(x_2\text{的系数},\ x_1\text{的系数})\\&=(12,5)\end{aligned}$$

因此, 最优对偶值计算如下:

$$\begin{aligned}(y_1,y_2)&=(x_2,x_1\text{的原始目标系数})\times(\text{最优逆矩阵})\\&=(12,5)\begin{pmatrix}\frac{2}{5} & -\frac{1}{5}\\ \frac{1}{5} & \frac{2}{5}\end{pmatrix}\\&=\left(\frac{29}{5},-\frac{2}{5}\right)\end{aligned}$$

原始–对偶目标值 在说明了如何确定最优对偶值后, 下面, 我们介绍原始目标值和对偶目标值之间的关系. 对于任意一对可行的原始解和对偶解, 都有:

$$\text{极大化问题的目标值}\leqslant\text{极小化问题的目标值}$$

在最优值处, 这个关系式作为严格等式成立. 这个关系式并没有特别指定哪一个问题是原始问题, 哪一个问题是对偶问题. 在这种情况下, 只有最优化含义 (极大化或极小化) 是重要的.

最优值不可能出现 z 严格小于 w(即 $z<w$), 因为无论 z 与 w 靠得如何近, 总是有余地可以改进, 这与图 4.2 演示的最优性矛盾.

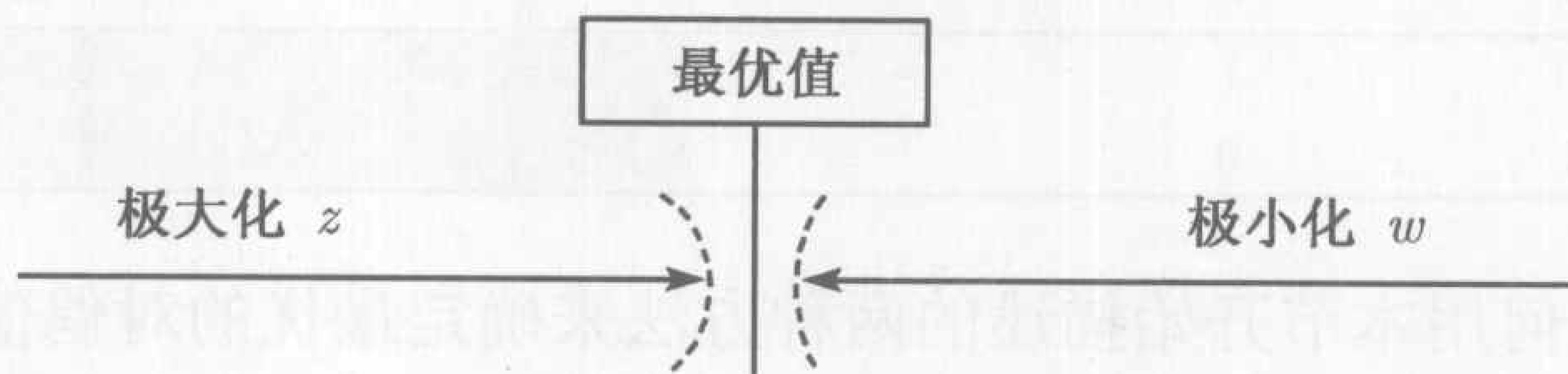


图 4.2　极大化 z 与极小化 w 之间的关系

例 4.2-2

在例 4.2-1 中, $\left(x_1=0, x_2=0, x_3=\frac{8}{3}\right)$ 和 $(y_1=6, y_2=0)$ 是可行的原始解与对偶解. 相应的目标值是

$$\begin{aligned} z &= 5x_1+12x_2+4x_3 = 5(0)+12(0)+4\left(\frac{8}{3}\right) = 10\frac{2}{3} \\ w &= 10y_1+8y_2 = 10(6)+8(0) = 60 \end{aligned}$$

因此, $z(=10\frac{2}{3})$ 对于极大化问题 (原始问题) 小于 $w(=60)$ 对于极小化问题 (对偶问题). z 的最优值 $(=54\frac{4}{5})$ 落在区间 $(10\frac{2}{3}, 60)$ 内.

习题 4.2C

1. 对下面的问题仅通过观察它的对偶问题来求目标函数的最优值. (不用单纯形法求解对偶问题.)

$$\begin{aligned} \min \quad & z = 10x_1+4x_2+5x_3 \\ \text{s.t.} \quad & 5x_1-7x_2+3x_3 \geqslant 50 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

2. 求解下面问题的对偶问题, 然后从对偶问题的解中求它的最优解. 对偶问题的解比直接求解原始问题提供了计算上的优势吗?

$$\begin{aligned} \min \quad & z = 5x_1+6x_2+3x_3 \\ \text{s.t.} \quad & 5x_1+5x_2+3x_3 \geqslant 50 \\ & x_1+x_2-x_3 \geqslant 20 \\ & 7x_1+6x_2-9x_3 \geqslant 30 \\ & 5x_1+5x_2+5x_3 \geqslant 35 \\ & 2x_1+4x_2-15x_3 \geqslant 10 \\ & 12x_1+10x_2 \geqslant 90 \\ & x_2-10x_3 \geqslant 20 \\ & x_1,\ x_2,\ x_3 \geqslant 0 \end{aligned}$$

*3. 考虑下面的线性规划:

$$\begin{aligned} \max\ z &= 5x_1 + 2x_2 + 3x_3 \\ \text{s.t.}\ \ x_1 + 5x_2 + 2x_3 &= 30 \\ x_1 - 5x_2 - 6x_3 &\leqslant 40 \\ x_1,\ x_2,\ x_3 &\geqslant 0 \end{aligned}$$

已知人工变量 x_4 和松弛变量 x_5 构成初始基变量, 在求解此问题时, 令 M 等于 100, 最优表给出如下:

基	x_1	x_2	x_3	x_4	x_5	解
z	0	23	7	105	0	150
x_1	1	5	2	1	0	30
x_5	0	−10	−8	−1	1	10

写出相应的对偶问题并使用两种方法确定它的最优解.

4. 考虑下面的线性规划:

$$\begin{aligned} \min\ \ z &= 4x_1 + x_2 \\ \text{s.t.}\ \ 3x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &\geqslant 6 \\ x_1 + 2x_2 &\leqslant 4 \\ x_1,\ x_2 &\geqslant 0 \end{aligned}$$

初始解由第 1 个、第 2 个约束的人工变量 x_4, x_5 与第 3 个约束的松弛变量 x_6 组成. 用 $M = 100$ 作为人工变量的系数, 最优表给出如下:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	0	0	0	−98.6	−100	−0.2	3.4
x_1	1	0	0	0.4	0	−0.2	0.4
x_2	0	1	0	0.2	0	0.6	1.8
x_3	0	0	1	1	−1	1	1.0

写出相应的对偶问题并使用两种方法确定它的最优解.

5. 考虑下面的线性规划:

$$\begin{aligned} \max\ z &= 2x_1 + 4x_2 + 4x_3 - 3x_4 \\ \text{s.t.}\ \ x_1 + x_2 + x_3 &= 4 \\ x_1 + 4x_2 + x_4 &= 8 \\ x_1,\ x_2,\ x_3,\ x_4 &\geqslant 0 \end{aligned}$$

用 x_3 和 x_4 作初始变量, 最优表给出如下:

基	x_1	x_2	x_3	x_4	解
z	2	0	0	3	16
x_3	0.75	0	1	−0.25	2
x_2	0.25	1	0	0.25	2

写出相应的对偶问题并使用两种方法确定它的最优解.

***6.** 考虑下面的线性规划:

$$\begin{aligned}\max\ & z = x_1 + 5x_2 + 3x_3 \\ \text{s.t.}\ & x_1 + 2x_2 + x_3 = 3 \\ & 2x_1 - x_2 \qquad = 4 \\ & x_1,\ x_2,\ x_3 \geqslant 0\end{aligned}$$

初始解由第 1 个约束的 x_3 和第 2 个约束的人工变量 x_4 组成, 人工变量的系数为 $M = 100$. 最优表给出如下:

基	x_1	x_2	x_3	x_4	解
z	0	2	0	99	5
x_3	0	2.5	1	−0.5	1
x_1	1	−0.5	0	0.5	2

写出相应的对偶问题并使用两种方法确定它的最优解.

7. 考虑下列不等式组:

$$\begin{aligned}2x_1 + 3x_2 &\leqslant 12 \\ -3x_1 + 2x_2 &\leqslant -4 \\ 3x_1 - 5x_2 &\leqslant 2 \\ x_1 &\text{无限制} \\ x_2 &\geqslant 0\end{aligned}$$

可以通过增加一个平凡的目标函数 $\max z = x_1 + x_2$ 求出可行解, 然后求解此问题. 另一种方法是求解对偶问题, 从对偶问题中可以求出该不等式组的一个解. 应用上述两种方法求解.

8. 估计下列线性规划最优目标函数值的区间:

***(a)**
$$\begin{aligned}\min\ & z = 5x_1 + 2x_2 \\ \text{s.t.}\ & x_1 - x_2 \geqslant 3 \\ & 2x_1 + 3x_2 \geqslant 5 \\ & x_1, x_2 \geqslant 0\end{aligned}$$

(b)
$$\begin{aligned}\max\ & z = x_1 + 5x_2 + 3x_3 \\ \text{s.t.}\ & x_1 + 2x_2 + x_3 = 3 \\ & 2x_1 - x_2 \qquad = 4 \\ & x_1, x_2, x_3 \geqslant 0\end{aligned}$$

(c)
$$\begin{aligned}\max\ & z = 2x_1 + x_2 \\ \text{s.t.}\ & x_1 - x_2 \leqslant 10 \\ & 2x_1 \qquad \leqslant 40 \\ & x_1, x_2 \geqslant 0\end{aligned}$$

(d)
$$\begin{aligned}\max\ & z = 3x_1 + 2x_2 \\ \text{s.t.}\ & 2x_1 + x_2 \leqslant 3 \\ & 3x_1 + 4x_2 \leqslant 12 \\ & x_1, x_2 \geqslant 0\end{aligned}$$

9. 在第 7(a) 题中, 令 y_1 和 y_2 是对偶变量. 确定下列原始–对偶解对是否最优.

***(a)** $(x_1 = 3, x_2 = 1; y_1 = 4, y_2 = 1)$ **(b)** $(x_1 = 4, x_2 = 1; y_1 = 1, y_2 = 0)$

(c) $(x_1 = 3, x_2 = 0; y_1 = 5, y_2 = 0)$

4.2.4 单纯形表的计算

这一节说明如何从问题的*原始*数据、与迭代相应的*逆矩阵*，以及对偶问题中产生出整个单纯形表的*任意一次*迭代. 使用图 4.1 单纯形表的示意图, 我们可以把计算分成两类:

(1) 约束列 (左端项和右端项); (2) 目标 z 行.

公式 1: 约束列的计算 在任意一次单纯形迭代中, 左端或右端项列的计算如下:

$$\text{第 } i \text{ 次迭代的约束列} = \text{第 } i \text{ 次迭代的逆矩阵} \times \text{原始的约束列}$$

公式 2: 目标 z 行计算 在任意一次单纯形迭代中, x_j 的目标方程系数 (简约费用) 的计算如下:

$$\text{原始 } z \text{ 方程中变量 } x_j \text{ 的系数} = \text{第 } j \text{ 个对偶约束的左端项} \\ - \text{第 } j \text{ 个对偶约束的右端项}$$

例 4.2-3

我们用例 4.2-2 的线性规划来解释公式 1 和公式 2 的应用. 从表 4.4 中的最优单纯形表, 我们有

$$\text{最优逆矩阵} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix}$$

通过计算最优表所有的左端项和右端项的列, 使用公式 1:

$$\begin{aligned} \text{最优迭代中的 } x_1 \text{ 列} &= \text{最优迭代中的逆矩阵} \times \text{原始的} x_1 \text{ 列} \\ &= \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

同理计算其余约束的列, 即有:

$$\text{最优迭代中的 } x_2 \text{ 列} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\text{最优迭代中的 } x_3 \text{ 列} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -\frac{1}{5} \\ \frac{7}{5} \end{pmatrix}$$

$$\text{最优迭代中的 } x_4 \text{ 列} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} \\ \frac{1}{5} \end{pmatrix}$$

$$\text{最优迭代中的 } R \text{ 列} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{5} \\ \frac{2}{5} \end{pmatrix}$$

$$\text{最优迭代中右端项的列} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \times \begin{pmatrix} 10 \\ 8 \end{pmatrix} = \begin{pmatrix} \frac{12}{5} \\ \frac{26}{5} \end{pmatrix}$$

接下来, 我们说明如何用公式 2 完成目标行的计算. 对偶变量的最优值, $(y_1, y_2) = (\frac{29}{5}, -\frac{2}{5})$, 例 4.2-2 中使用了两种不同的方法给出了相应的计算结果. 这些值在公式 2 中用于确定相应的 z 系数, 即有:

$$\begin{aligned}
x_1 \text{ 的 } z \text{ 系数} &= y_1 + 2y_2 - 5 &&= \tfrac{29}{5} + 2 \times \left(-\tfrac{2}{5}\right) - 5 &&= 0 \\
x_2 \text{ 的 } z \text{ 系数} &= 2y_1 - y_2 - 12 &&= 2 \times \tfrac{29}{5} - \left(-\tfrac{2}{5}\right) - 12 &&= 0 \\
x_3 \text{ 的 } z \text{ 系数} &= y_1 + 3y_2 - 4 &&= \tfrac{29}{5} + 3 \times \left(-\tfrac{2}{5}\right) - 4 &&= \tfrac{3}{5} \\
x_4 \text{ 的 } z \text{ 系数} &= y_1 - 0 &&= \tfrac{29}{5} - 0 &&= \tfrac{29}{5} \\
R \text{ 的 } z \text{ 系数} &= y_2 - (-M) &&= -\tfrac{2}{5} - (-M) &&= -\tfrac{2}{5} + M
\end{aligned}$$

注意, 公式 1 和公式 2 的计算能够用在原始问题或者对偶问题的任意一步迭代中. 这里我们只需要相应于 (原始或对偶) 迭代的逆矩阵和原始线性规划数据.

习题 4.2D

1. 生成例 4.2-1 的第 1 个单纯形迭代 (为方便起见, 可以使用 TORA 的 Iterations ⇒ M-method), 然后用公式 1 和公式 2 验证所得表中的全部元素.
2. 考虑下面的线性规划模型:

$$\begin{aligned}
\max \quad & z = 4x_1 + 14x_2 \\
\text{s.t.} \quad & 2x_1 + 7x_2 + x_3 = 21 \\
& 7x_1 + 2x_2 + x_4 = 21 \\
& x_1, x_2, x_3, x_4 \geqslant 0
\end{aligned}$$

检验下列每一个基本解的最优性与可行性.

*(a) 基变量 $= (x_2, x_4)$, 逆矩阵 $= \begin{pmatrix} \frac{1}{7} & 0 \\ -\frac{2}{7} & 1 \end{pmatrix}$

(b) 基变量 $= (x_2, x_3)$, 逆矩阵 $= \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{7}{2} \end{pmatrix}$

(c) 基变量 $= (x_2, x_1)$, 逆矩阵 $= \begin{pmatrix} \frac{7}{45} & -\frac{2}{45} \\ -\frac{2}{45} & \frac{7}{45} \end{pmatrix}$

(d) 基变量 $= (x_1, x_4)$, 逆矩阵 $= \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{7}{2} & 1 \end{pmatrix}$

3. 考虑下面的线性规划模型:

$$
\begin{aligned}
\max\quad & z = 3x_1 + 2x_2 + 5x_3 \\
\text{s.t.}\quad & x_1 + 2x_2 + x_3 + x_4 = 30 \\
& 3x_1 + 2x_3 + x_5 = 60 \\
& x_1 + 4x_2 + x_6 = 20 \\
& x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6 \geqslant 0
\end{aligned}
$$

检验下列每一个基本解的最优性与可行性.

(a) 基变量 $= (x_4, x_3, x_6)$, 逆矩阵 $= \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(b) 基变量 $= (x_2, x_3, x_1)$, 逆矩阵 $= \begin{pmatrix} \frac{1}{4} & -\frac{1}{8} & \frac{1}{8} \\ \frac{3}{2} & -\frac{1}{4} & -\frac{3}{4} \\ -1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

(c) 基变量 $= (x_2, x_3, x_6)$, 逆矩阵 $= \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix}$

***4.** 考虑下面的线性规划模型:

$$
\begin{aligned}
\min\quad & z = 2x_1 + x_2 \\
\text{s.t.}\quad & 3x_1 + x_2 - x_3 = 3 \\
& 4x_1 + 3x_2 - x_4 = 6 \\
& x_1 + 2x_2 + x_5 = 3 \\
& x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geqslant 0
\end{aligned}
$$

对应于下面的基本解计算出完整的单纯形表, 并检验它的最优性和可行性.

$$
\text{基变量} = (x_1, x_2, x_5),\ \text{逆矩阵} = \begin{pmatrix} \frac{3}{5} & -\frac{1}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \\ 1 & -1 & 1 \end{pmatrix}
$$

5. 考虑下面的线性规划模型:

$$
\begin{aligned}
\max\quad & z = 5x_1 + 12x_2 + 4x_3 \\
\text{s.t.}\quad & x_1 + 2x_2 + x_3 + x_4 = 10 \\
& 2x_1 - x_2 + 3x_3 = 2 \\
& x_1,\ x_2,\ x_3,\ x_4 \geqslant 0
\end{aligned}
$$

(a) 从下列基本可行解中找出最好的解.

(i) 基变量 $= (x_4, x_3)$, 逆矩阵 $= \begin{pmatrix} 1 & -\frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$

(ii) 基变量 $= (x_2, x_1)$, 逆矩阵 $= \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix}$

(iii) 基变量 $= (x_2, x_3)$, 逆矩阵 $= \begin{pmatrix} \frac{3}{7} & -\frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix}$

(b) 对于线性规划模型, 由 (a) 得到的解是最优的吗?

6. 考虑下面的线性规划模型:

$$\begin{aligned} \max \quad & z = 5x_1 + 2x_2 + 3x_3 \\ \text{s.t.} \quad & x_1 + 5x_2 + 2x_3 \leqslant b_1 \\ & x_1 - 5x_2 - 6x_3 \leqslant b_2 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

下表是对应于 b_1 和 b_2 特定值的最优表:

基	x_1	x_2	x_3	x_4	x_5	解
z	0	a	7	d	e	150
x_1	1	b	2	1	0	30
x_5	0	c	-8	-1	1	10

求以下各值:

(a) 右端项 b_1 和 b_2 的值. (b) 最优对偶解. (c) 元素 a, b, c, d, e.

***7.** 下表是带有 3 个 ($\leqslant$) 约束和所有非负变量的一个求极大值的线性规划模型的最优单纯形表. 变量 x_3, x_4, x_5 为相应于 3 个约束的松弛变量. 使用原始和对偶目标函数采用不同的方法求出相应的最优目标值.

基	x_1	x_2	x_3	x_4	x_5	解
z	0	0	0	3	2	?
x_3	0	0	1	1	-1	2
x_2	0	1	0	1	0	6
x_1	1	0	0	-1	1	2

8. 考虑下面的线性规划:

$$\begin{aligned} \max \quad & z = 2x_1 + 4x_2 + 4x_3 - 3x_4 \\ \text{s.t.} \quad & x_1 + x_2 + x_3 = 4 \\ & x_1 + 4x_2 + x_4 = 8 \\ & x_1, x_2, x_3, x_4 \geqslant 0 \end{aligned}$$

用对偶问题证明基本解 (x_1, x_2) 不是最优的.

9. 证明在 4.2.3 节中确定最优对偶值的方法 1 实际上是可由 4.2.4 节中的公式 2 推导出来.

4.3 对偶的经济学解释

线性规划问题可以看成是一种资源分配模型, 模型中的目标是极大化收入, 约束是有限资源的利用. 从这个观点看问题, 相应的对偶问题提供了线性规划资源分配模型有趣的经济学解释.

使用公式进行讨论, 我们考虑下面一般原始问题与对偶问题的表示:

原始问题	对偶问题
$\max \quad z=\sum_{j=1}^{n} c_j x_j$	$\min \quad w=\sum_{i=1}^{m} b_i y_i$
s.t. $\quad \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \quad i=1,2,\cdots,m$	s.t. $\quad \sum_{i=1}^{m} a_{ij} y_i \geqslant c_j, \quad j=1,2,\cdots,n$
$x_j \geqslant 0, \quad j=1,2,\cdots,n$	$y_i \geqslant 0, \quad i=1,2,\cdots,m$

将问题看成是资源分配模型, 原始问题有 n 种经济活动和 m 种资源. 在原始问题中系数 c_j 表示活动 j 的单位收入. 资源 i 的最大可用量为 b_i, 且资源 i 在活动 j 中的单位消耗率是 a_{ij}.

4.3.1 对偶变量的经济学解释

4.2.3 节曾说明, 对任意原始可行解和对偶可行解, 当目标函数值有限时, 一定满足下面的不等式:

$$z=\sum_{j=1}^{n} c_j x_j \leqslant \sum_{i=1}^{m} b_i y_i = w$$

当原始解与对偶解是最优解时, 等式严格成立, 即 $z=w$.

首先, 考虑最优条件 $z=w$. 已知原始问题表示资源分配模型, 我们能够认为 z 表示收入的美元数. 因为 b_i 表示资源 i 可用的单位数, 方程 $z=w$ 可表示成

$$\$=\sum_{i} \text{资源 } i \text{ 的单位数} \times \text{每单位资源的 } \$$$

这意味着对偶变量 y_i 表示资源 i 的**单位价值**. 如 3.6 节所述, 标准的名称是资源 i 的**对偶价格**或**影子价格**, 在所有线性规划的文献和软件包中都用它来替代名称单位价值.

按照相同的逻辑, 不等式 $z<w$, 可以将相应的原始可行解与对偶可行解解释为

$$\text{收入} < \text{资源的价值}$$

这个关系式说明, 只要从全部活动中得到的总收入小于资源的价值, 相应的原始解和对偶解就不是最优的. 只有当资源完全地被开发, 最优性 (最大收入) 才能达到, 这只能发生在输入 (资源的价值) 等于输出 (收入的美元数) 时. 从经济学的角度来

看, 当输入 (资源的价值) 超过输出 (收入), 系统被称为不稳定的(非最优的). 稳定仅出现在两个量相等时.

例 4.3-1

已知 Reddy Mikks 模型 (例 2.1-1) 和它的对偶问题如下:

Reddy Mikks 原始问题	Reddy Mikks 对偶问题
max $z = 5x_1 + 4x_2$	min $w = 24y_1 + 6y_2 + y_3 + 2y_4$
s.t. $6x_1 + 4x_2 \leqslant 24$ (资源 1, M1)	s.t. $6y_1 + y_2 - y_3 \geqslant 5$
$x_1 + 2x_2 \leqslant 6$ (资源 2, M2)	$4y_1 + 2y_2 + y_3 + y_4 \geqslant 4$
$-x_1 + x_2 \leqslant 1$ (资源 3, 市场)	$y_1, y_2, y_3, y_4 \geqslant 0$
$x_2 \leqslant 2$ (资源 4, 需求)	
$x_1, x_2 \geqslant 0$	
最优解:	最优解:
$x_1 = 3, x_2 = 1.5, z = 21$	$y_1 = 0.75, y_2 = 0.5, y_3 = y_4 = 0,$ $w = 21$

简要地说, Reddy Mikks 模型涉及生产两种类型的涂料 (内墙和外墙), 这两种产品使用两种原料 M1 和 M2(资源 1 和资源 2) 并受市场和需求的限制, 由第 3 个和第 4 个约束来表示. 模型确定内墙和外墙涂料的数量 (吨/天), 使得每天的收入最大化 (单位是 1 000 美元).

最优对偶解说明原料 1(资源 1) 的对偶价格 (单位价值) 是 $y_1 = 0.75$(或每吨 750 美元), 原料 2(资源 2) 的对偶价格是 $y_2 = 0.5$(或每吨 500 美元). 这些结果对于特定的、正如我们在 3.6 节说明的可行性区域也适用. 对于表示市场和需求限制的资源 3 和资源 4, 对偶价格均为零, 这表明它们相应的资源是充裕的. 因此, 它们的单位价值是零.

习题 4.3A

1. 在例 4.3-1 中, 计算下面每种情况下最优收入的变化 (用 TORA 输出得到可行性区域).

(a) 原料 M1(资源 1) 的约束是 $6x_1 + 4x_2 \leqslant 22$.

(b) 原料 M2(资源 2) 的约束是 $x_1 + 2x_2 \leqslant 4.5$.

(c) 由资源 4 表示的市场条件是 $x_2 \leqslant 10$.

***2.** NWAC 电力公司为军事承包商生产 4 种类型的简单电缆. 每种电缆必须经过 4 种相继的操作: 拼接、焊接、套管和检查. 下表给出了该问题相关的数据.

电缆	单位分钟数				单位收入 ($)
	拼接	焊接	套管	检查	
SC320	10.5	20.4	3.2	5.0	9.40
SC325	9.3	24.6	2.5	5.0	10.80
SC340	11.6	17.7	3.6	5.0	8.75
SC370	8.2	26.5	5.5	5.0	7.80
日能力 (分钟)	4 800.0	9 600.0	4 700.0	4 500.0	

承包商保证对于 4 种电缆的每一种最低产量是 100 个单位.

(a) 将问题建立成一个线性规划模型, 并确定最优的生产规划.

(b) 根据对偶价格, 你会推荐增加 4 种操作中任何一种的日操作能力吗? 试解释.

(c) 对于 4 种电缆的最低产量要求对 NWAC 电力公司有利还是不利? 试根据对偶价格给出解释.

(d) 如果将焊接能力增加 10%, 能够保证由对偶价格指定的现有单位收入吗?

3. BagCo 公司生产皮夹克和手提包. 一件皮夹克需要 8 m^2 的皮革, 一个手提包仅需要 2 m^2 的皮革. 对于两种产品的劳动力需求分别是 12 小时和 5 小时. 目前一周提供的皮革和劳动力的限制是 1 200 m^2 和 1 850 小时. 公司销售皮夹克和手提包的价格分别是 350 美元和 120 美元. 公司的目标是确定生产规划使净收入为最大. 如果 BagCo 公司考虑扩大生产, 那么公司购买额外皮革需要支付的最高价格是多少? 为额外劳动力需要支付的工资是多少?

4.3.2 对偶约束的经济学解释

对偶约束可以用 4.2.4 节中的公式 2 来解释, 它说明在任何原始问题的迭代中,

$$\begin{aligned} x_j\text{的目标系数} &= \text{对偶约束 } j\text{的左端项} - \text{对偶约束 } j\text{的右端项} \\ &= \sum_{i=1}^{m} a_{ij}y_i - c_j \end{aligned}$$

我们再次用维数分析来解释这个方程. 活动 j 的单位收入 c_j 以美元计. 因此, 从一致性的角度来看, $\sum_{i=1}^{m} a_{ij}y_i$ 的量纲也一定是美元/单位. 其次, 因为 c_j 表示收入, $\sum_{i=1}^{m} a_{ij}y_i$ 以相反的符号出现在方程中, 一定表示成本. 因此, 我们有

$$\$\text{成本} = \sum_{i=1}^{m} a_{ij}y_i = \sum_{i=1}^{m} \text{每单位活动 } j \text{ 对应于资源 } i \text{ 的用量} \times \text{资源 } i \text{ 的单位成本}$$

这里的结论是对偶变量 y_i 表示每单位资源 i 的**估算成本** (imputed cost), 因此我们能够把量 $\sum_{i=1}^{m} a_{ij}y_i$ 看作是产生一个单位活动 j 所需全部资源的估算成本.

在 3.6 节中, 我们提到量 $\left(\sum_{i=1}^{m} a_{ij}y_i - c_j\right)$ 作为活动 j 的**简约费用**. 单纯形法的极大化问题的最优性条件是说, 仅当无用 (非基) 活动 j 的简约费用为负, 对它的投入增加才能够提高收入. 按照前面的解释, 这个条件说明

$$\text{用在一个单位活动 } j \text{ 上的资源的估算成本} < \text{活动 } j \text{ 的单位收入}$$

因此, 极大化问题的最优性条件是说, 如果一项活动的单位收入超过它的单位估算成本, 那么从经济学角度来看, 增加该项活动的投入就能够增加收入.

我们将用 3.6 节的 TOYCO 模型演示其计算. 为了方便, 这里再陈述一遍模型的细节.

例 4.3-2

TOYCO 公司用 3 种操作装配 3 种玩具：玩具火车、玩具卡车和玩具汽车. 对于 3 种操作的可利用的装配时间分别是每天 430 分钟、460 分钟和 420 分钟, 每个玩具火车、玩具卡车和玩具汽车的收入分别是 3 美元、2 美元和 5 美元. 每辆火车在 3 种操作的装配时间花费分别是 1 分钟、3 分钟和 1 分钟. 每辆卡车和每辆汽车相应的时间是 $(2,0,4)$ 和 $(1,2,0)$ 分钟 (零时间表示未用到此项操作).

令 x_1, x_2, x_3 表示玩具火车、玩具卡车和玩具汽车每天的装配单位数量, 相应的线性规划模型和它的对偶给出如下：

TOYCO 原始问题	TOYCO 对偶问题
max $z = 3x_1 + 2x_2 + 5x_3$	min $w = 430y_1 + 460y_2 + 420y_3$
s.t. $x_1 + 2x_2 + x_3 \leqslant 430$ (操作 1)	s.t. $y_1 + 3y_2 + y_3 \geqslant 3$
$3x_1 + 2x_3 \leqslant 460$ (操作 2)	$2y_1 + 4y_3 \geqslant 2$
$x_1 + 4x_2 \leqslant 420$ (操作 3)	$y_1 + 2y_2 \geqslant 5$
$x_1, x_2, x_3 \geqslant 0$	$y_1, y_2, y_3 \geqslant 0$
最优解：	最优解：
$x_1 = 0, x_2 = 100, x_3 = 230, z = \$1\ 350$	$y_1 = 1, y_2 = 2, y_3 = 0, w = \$1\ 350$

最优原始解要求不生产玩具火车, 生产 100 辆玩具卡车和 230 辆玩具汽车. 假定 TOYCO 对生产玩具火车也感兴趣. 将如何完成这件事呢? 用 x_1 的简约费用来解释这个问题, 玩具火车变成有经济吸引力只有当用于生产一辆玩具火车的资源估算成本小于它的单位收入. 因此, TOYCO 可以通过提高单位价格的方法来增加每辆玩具火车的单位收入, 或者它减少所用资源的估算成本 $(= y_1 + 3y_2 + y_3)$. 增加单位价格, 由于市场竞争, 可能并不可行. 减少单位估算成本看起来更有道理, 因为它需要改进装配操作. 令 r_1, r_2, r_3 为 3 种操作单位时间的减少比例, 问题是需要确定 r_1, r_2, r_3, 使得每辆玩具火车新的估算成本小于它的单位收入, 也就是,

$$1(1 - r_1)y_1 + 3(1 - r_2)y_2 + 1(1 - r_3)y_3 < 3$$

对于给出的最优值 $y_1 = 1, y_2 = 2$ 和 $y_3 = 0$, 这个不等式简化为 (请验证!)

$$r_1 + 6r_2 > 4$$

因此, 任意的在 0 到 1 之间且满足 $r_1 + 6r_2 > 4$ 的 r_1 和 r_2 能够使玩具火车有利可图. 然而, 这个目标或许并不能实现, 因为它实际需要的操作 1 和操作 2 的时间要减少是不可能的. 例如, 即使在这两项操作的时间上减少的比例高达 50%(也就是, $r_1 = r_2 = 0.5$), 对于给定的条件还是失败的, 因此, TOYCO 将不再生产玩具火车, 除非有增加单位价格的可能性.

习题 4.3B

1. 在例 4.3-2 中, 假定对于玩具火车, 操作 2 的单位时间从 3 分钟最多减少到 1.25 分钟. 操作 1 的单位时间必须减少到多少, 才使生产玩具火车恰好有利可图?

***2.** 在例 4.3-2 中, 假定 TOYCO 正在研究引进第 4 种玩具的可能性: 玩具救火车. 其装配不使用操作 1, 且操作 2 和操作 3 的单位装配时间分别是 1 分钟和 3 分钟. 每辆玩具救火车的收入是 4 美元. 你是否建议 TOYCO 引进这种新产品?

***3.** JoShop 公司使用车床和钻压生产 4 种类型的机器零件 PP1、PP2、PP3 和 PP4. 下表总括了相关数据.

机器	加工单个零件的时间数				能力 (分钟)
	PP1	PP2	PP3	PP4	
车床	2	5	3	4	5 300
钻压	3	4	6	4	5 300
单位收入 ($)	3	6	5	4	

对于当前最优解不生产的零件, 求每增加一个这些产品所引起的最优收入的下降率.

4. 考虑在第 3 题中 JoShop 公司的最优解. 公司估计, 对于不生产的零件 (由最优解确定) 通过生产过程的改进, 可全面减少 20% 的机器时间. 这项改进能使这些零件有利可图吗? 如果不能, 需要下降的最小百分比是多少才能有收益?

4.4 其他单纯形算法

在第 3 章提到的单纯形算法中, 问题从一个 (基本) 可行解开始. 后面的迭代保持可行性, 直到最后一步迭代达到最优解. 这种算法有时称为**原始单纯形** (primal simplex) 法.

本节介绍其他两种算法: **对偶单纯形** (dual simplex) 法和**广义单纯形** (generalized simplex) 法. 在对偶单纯形法中, 线性规划从一个优于最优解的不可行(基本)解开始. 以后的迭代仍保持不可行并 (优于) 最优, 直到最后一次迭代恢复可行性. 广义单纯形法将原始单纯形法和对偶单纯形法组合成一个算法. 它处理在开始处既不是最优又不是可行的解的问题. 在这个算法中, 相继的迭代由基本可行或不可行 (基本) 解构成. 在最终的迭代中, 解变为最优且可行的 (假定它们存在).

用于后最优分析计算过程中的所有 3 种算法 (原始单纯形法、对偶单纯形法和广义单纯形法) 将在 4.5 节中介绍.

4.4.1 对偶单纯形算法

对偶单纯形法的核心是以一个优于最优解且不可行的基本解开始, 最优性条件和可行性条件用于在保持基本解的最优性的同时使解的迭代朝着可行性移动.

对偶可行性条件 离基变量 x_r 是具有最负值 (多个时任选其一) 的基变量. 如果所有的基变量均是非负的, 算法结束.

对偶最优性条件 已知 x_r 是离基变量, 令 $\bar{c}_j$ 是非基变量 x_j 的简约费用, α_{rj} 是单纯形表中约束在 x_r 行和 x_j 列的系数. 进基变量是非基变量, 其中 $\alpha_{rj} < 0$ 且使得下式有相应的最小比

$$\min_{\text{非基}x_j} \left\{ \left| \frac{\bar{c}_j}{\alpha_{rj}} \right|, \alpha_{rj} < 0 \right\}$$

(多个时任选其一.) 如果 $\alpha_{rj} \geqslant 0$ 对于所有的非基变量 x_j 成立, 则问题没有可行解.

从一个线性规划最优且不可行点开始, 下面两个条件必须得到满足:

(1) 目标函数满足一般单纯形法的最优性条件 (第 3 章).

(2) 所有约束的类型是 $(\leqslant)$.

第二个条件需要将任何 $(\geqslant)$ 转换为 $(\leqslant)$, 可以通过简单地在不等式 $(\geqslant)$ 的两边乘上 -1 来转换. 若线性规划包括 $(=)$ 约束, 则等式约束可以用两个不等式来代替. 例如,

$$x_1 + x_2 = 1$$

等价于

$$x_1 + x_2 \leqslant 1, \quad x_1 + x_2 \geqslant 1$$

或

$$x_1 + x_2 \leqslant 1, \quad -x_1 - x_2 \leqslant -1$$

在把所有约束转换成 $(\leqslant)$ 后, 如果不等式右端项中至少有一个是严格的负数, 那么初始解是不可行的.

例 4.4-1

$$\begin{aligned} \min \quad & z = 3x_1 + 2x_2 + x_3 \\ \text{s.t.} \quad & 3x_1 + x_2 + x_3 \geqslant 3 \\ & -3x_1 + 3x_2 + x_3 \geqslant 6 \\ & x_1 + x_2 + x_3 \leqslant 3 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

在当前的例子中, 将第 1 个和第 2 个不等式乘上 -1, 转换成 $(\leqslant)$ 约束. 因此, 初始单纯形表将如下给出:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	-3	-2	-1	0	0	0	0
x_4	-3	-1	-1	1	0	0	-3
x_5	3	-3	-1	0	1	0	-6
x_6	1	1	1	0	0	1	3

这个表是最优的, 因为 z 行中的所有简约费用均 $\leqslant 0$ ($\bar{c}_1=-3$, $\bar{c}_2=-2$, $\bar{c}_3=-1$, $\bar{c}_4=0$, $\bar{c}_5=0$, $\bar{c}_6=0$). 它同时还是不可行的, 因为基变量中至少有一个取值为负 ($x_4=-3$, $x_5=-6$, $x_6=3$).

按照对偶可行性条件, $x_5(=-6)$ 是离基变量. 下表显示如何用对偶最优性条件确定进基变量.

	$j=1$	$j=2$	$j=3$
非基变量	x_1	x_2	x_3
z 行 ($\bar{c}_j$)	-3	-2	-1
x_5 行, α_{5j}	3	-3	-1
比值, $\left\lvert\frac{\bar{c}_j}{\alpha_{5j}}\right\rvert$, $\alpha_{5j}<0$	—	$\frac{2}{3}$	1

下表可用我们熟悉的行运算得到, 给出如下:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	-5	0	$-\frac{1}{3}$	0	$-\frac{2}{3}$	0	4
x_4	-4	0	$-\frac{2}{3}$	1	$-\frac{1}{3}$	0	-1
x_2	-1	1	$\frac{1}{3}$	0	$-\frac{1}{3}$	0	2
x_6	2	0	$\frac{2}{3}$	0	$\frac{1}{3}$	1	1
比值	$\frac{5}{4}$	—	$\frac{1}{2}$	—	2	—	

上表说明 x_4 离基且 x_3 进基, 于是可得下表, 它既是最优的又是可行的:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	-3	0	0	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{9}{2}$
x_3	6	0	1	$-\frac{3}{2}$	$\frac{1}{2}$	0	$\frac{3}{2}$
x_2	-3	1	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{3}{2}$
x_6	-2	0	0	1	0	1	0

注意对偶单纯形的工作原理. 在所有的迭代中, 一直都要保持最优性 (所有简约费用 $\leqslant 0$). 同时, 每一个新的迭代将解向可行性方向移动. 在迭代 3 中, 第一次恢复可行性, 因此迭代过程结束, 并得到最优的可行解 $x_1=0$, $x_2=\frac{3}{2}$, $x_3=\frac{3}{2}$ 和 $z=\frac{9}{2}$.

TORA 程序

TORA 为对偶单纯形方法提供了一教学模块. 从 SOLVE/MODIFY 菜单中

选择 Solve ⇒ Algebraic ⇒ Iterations ⇒ Dual Simplex. 记住, 你需要将 (=) 约束转换成不等式约束. 你并不需要转换 (⩾) 约束, 因为 TORA 将在内部作转换. 如果线性规划并不满足对偶单纯形法的初始必要条件, 提示信息将出现在屏幕上.

像在正规的单纯形方法中一样, 该教学模块允许你预先选择进基和离基变量. 然后, 它会给你适当的反馈, 告诉你选择的是否正确.

习题 4.4A ①

1. 考虑图 4.3 中的解空间, 要求用对偶单纯形法对 $\min z = 2x_1 + x_2$ 求出最优极点. 最优解出现在图中的点 $F = (0.5, 1.5)$ 处.

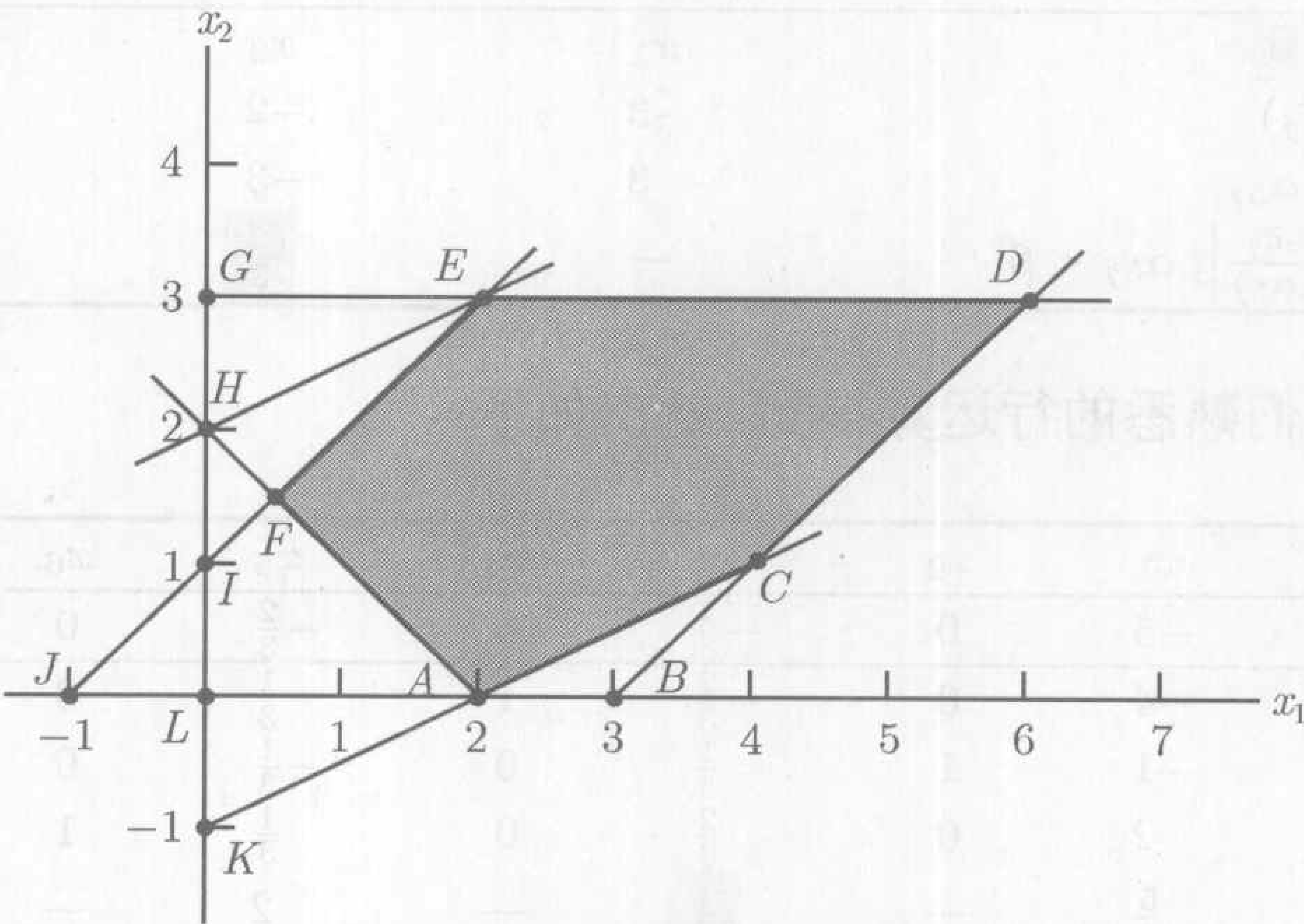


图 4.3　习题 4.4A 第 1 题的解空间

(a) 对偶单纯形法能从点 A 开始吗?

***(b)** 如果初始基本解 (不可行但好于最优点) 由点 G 给出, 对偶单纯形法以路径 $G \to E \to F$ 的迭代是否可能? 试解释.

(c) 如果初始基本 (不可行) 解在点 L 处开始, 确定一条到达最优可行点 F、由对偶单纯形法产生的可能路径.

2. 为下面的问题产生对偶单纯形迭代 (为方便起见, 使用 TORA), 并跟踪算法在图解空间上的路径.

(a)
$$\begin{aligned} \min \quad & z = 2x_1 + 3x_2 \\ \text{s.t.} \quad & 2x_1 + 2x_2 \leqslant 30 \\ & x_1 + 2x_2 \geqslant 10 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

(b)
$$\begin{aligned} \min \quad & z = 5x_1 + 6x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geqslant 2 \\ & 4x_1 + x_2 \geqslant 4 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

---

① 鼓励你使用 TORA 的教学模块, 这里可能避免完成高斯–若尔当行运算单调乏味的任务. 照这样, 你能够集中在理解方法的主要概念.

(c) $\min\ z = 4x_1 + 2x_2$

s.t. $x_1 + x_2 = 1$

$3x_1 - x_2 \geqslant 2$

$x_1, x_2 \geqslant 0$

(d) $\min\ z = 2x_1 + 3x_2$

s.t. $2x_1 + x_2 \geqslant 3$

$x_1 + x_2 = 2$

$x_1, x_2 \geqslant 0$

3. 带有人工约束的对偶单纯形法. 考虑下面的问题:

$$\begin{aligned} \max\ & z = 2x_1 - x_2 + x_3 \\ \text{s.t.}\ & 2x_1 + 3x_2 - 5x_3 \geqslant 4 \\ & -x_1 + 9x_2 - x_3 \geqslant 3 \\ & 4x_1 + 6x_2 + 3x_3 \leqslant 8 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

初始基本解由剩余变量 x_4, x_5 以及松弛变量 x_6 组成, 它是不可行的, 因为 $x_4 = -4$, $x_5 = -3$. 然而, 对偶单纯形并不能直接应用, 因为 x_1 和 x_3 不满足极大化问题的最优性条件. 证明: 增加人工约束 $x_1 + x_3 \leqslant M$(这里 M 充分大, 不能消去原解空间中任何可行点), 然后用新的约束作为枢轴行, 选择 x_1 作为进基变量 (因为它有最负的目标系数) 将给出所有的最优目标行. 接下来, 在所修正的问题上完成正规的对偶单纯形法.

4. 用第 3 题引进的人工约束过程, 用对偶单纯形法求解下面的问题. 对于每种情况, 指出得到的解是可行、不可行或无界.

(a) $\max\ z = 2x_3$

s.t. $-x_1 + 2x_2 - 2x_3 \geqslant 8$

$-x_1 + x_2 + x_3 \leqslant 4$

$2x_1 - x_2 + 4x_3 \leqslant 10$

$x_1, x_2, x_3 \geqslant 0$

(b) $\max\ z = x_1 - 3x_2$

s.t. $x_1 - x_2 \leqslant 2$

$x_1 + x_2 \geqslant 4$

$2x_1 - 2x_2 \geqslant 3$

$x_1, x_2 \geqslant 0$

*(c) $\min\ z = -x_1 + x_2$

s.t. $x_1 - 4x_2 \geqslant 5$

$x_1 - 3x_2 \leqslant 1$

$2x_1 - 5x_2 \geqslant 1$

$x_1, x_2 \geqslant 0$

(d) $\max\ z = 2x_3$

s.t. $-x_1 + 3x_2 - 7x_3 \geqslant 5$

$-x_1 + x_2 - x_3 \leqslant 1$

$3x_1 + x_2 - 10x_3 \leqslant 8$

$x_1, x_2, x_3 \geqslant 0$

5. 使用 3 种不同的方法求解下面的线性规划问题 (为方便, 使用 TORA). 哪种方法显示出最有效的计算?

$$\begin{aligned} \min\ & z = 6x_1 + 7x_2 + 3x_3 + 5x_4 \\ \text{s.t.}\ & 5x_1 + 6x_2 - 3x_3 + 4x_4 \geqslant 12 \\ & x_2 - 5x_3 - 6x_4 \geqslant 10 \\ & 2x_1 + 5x_2 + x_3 + x_4 \geqslant 8 \\ & x_1, x_2, x_3, x_4 \geqslant 0 \end{aligned}$$

4.4.2 广义单纯形算法

第 3 章介绍的 (原始) 单纯形算法以一个可行但不是最优的解开始. 4.4.1 节介绍的对偶单纯形法以一个 (优于) 最优值但不可行的解开始. 如果线性规划模型的

开始点既不是最优的又不是可行的, 将如何做呢? 我们已经看到, 原始单纯形法借助于人工变量解决初始解的不可行性. 类似地, 对偶单纯形法使用人工约束来解决非最优性 (见习题 4.4A 的第 3 题). 尽管设计这些过程用于提高自动计算能力, 但这些细节可能让人们忘记单纯形算法最本质的东西 —— 线性规划的最优解与某个角点 (或基本) 解相对应. 基于这个事实, 对于从非最优和非可行点开始的线性规划模型, 你应该能够 "量身定制" 自己的单纯形算法. 下面的例子解释了所谓的广义单纯形算法.

例 4.4-2

考虑习题 4.4A 中第 4 题的线性规划模型. 模型可以构造成下列单纯形表形式, 其初始解 (x_4, x_5, x_6) 既是非最优的 (因为 x_3 有负的简约费用) 又是不可行的 (因为 $x_4 = -8$). (第一个方程乘上 -1, 直接展示了解列的不可行性.)

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	0	0	-2	0	0	0	0
x_4	1	-2	2	1	0	0	-8
x_5	-1	1	1	0	1	0	4
x_6	2	-1	4	0	0	1	10

可以在不使用任何人工变量和人工约束的情况下求解此问题, 其方法如下: 首先通过应用对偶单纯形可行性条件消除不可行性, 选择 x_4 作为离基变量. 为确定进基变量, 我们所需要的只是一个非基变量, 它的约束系数在 x_4 行是严格负的. 这种选择可以不考虑最优性, 因为在此点最优性是不存在的 (与对偶最优性条件比较). 在当前的例子中, x_2 在 x_4 行中有一个负的系数, 选择 x_2 作为进入变量. 其结果如下表所示:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	0	0	-2	0	0	0	0
x_2	$-\frac{1}{2}$	1	-1	$-\frac{1}{2}$	0	0	4
x_5	$-\frac{1}{2}$	0	2	$\frac{1}{2}$	1	0	0
x_6	$\frac{3}{2}$	0	3	$-\frac{1}{2}$	0	1	14

上表中的解目前是可行的, 但不是最优的, 因此我们可以使用原始单纯形法确定出最优解. 在一般情况下, 假如还不能恢复上表的可行性, 我们还需要重复上述过程直到可行性满足为止, 或者有证据表明问题没有可行解 (如果有一个基变量是负的, 而所有约束的系数是非负的, 会发生这种情况). 一旦建立了可行性, 下一步是借助于原始单纯形法适当的最优性条件, 将注意力转向最优性方面.

评注 例 4.4-2 的本质是, 单纯形法并非死板的, 有大量的文献介绍单纯形法的各

种变形 [例如, 原始–对偶方法、对称方法、交叉 (criss-cross) 方法和多元方法]. 它们给人的印象是, 虽然每种方法都是不同的, 但实际上它们都在寻找角点解, 并且都在朝着自动计算, 或许还有计算的效率方面努力.

习题 4.4B

1. 习题 4.4A 第 4(c) 题的线性规划模型没有可行解. 说明这个条件如何用广义单纯形过程来检验.
2. 习题 4.4A 第 4(d) 题的线性规划模型没有有界解. 说明这个条件如何用广义单纯形过程来检验.

4.5 后最优分析

3.6 节通过确定会使最优基本解保持不变的不同参数的变化区间, 介绍了最优解的灵敏度分析. 接下来的一个自然的问题就是*后最优分析*, 其目的是让模型参数发生一定的改变并求出新的最优解. 我们举一个关于家禽业的例子, 这里我们通常用线性规划模型来确定每只雏鸡的最优饲料配制 (见例 2.2-2). 每只雏鸡每周的饲料消耗量从 0.26 磅 (120 克, 喂养一周大的雏鸡) 到 2.1 磅 (950 克, 喂养 8 周大的嫩鸡). 此外, 混合饲料中成分的成本可能会定期变化. 这些变化需要定期重新计算最优解. *后最优分析*采用有效的方法确定新的最优解. 新的最优解用 4.2 节介绍的对偶性和原始–对偶关系来求出.

下表给出了后最优分析中的各种情况, 以及为得到新的最优解所需要采取的行动 (假设存在一种情况):

参数改变后的条件	推荐采用的行动
当前解保持最优与可行	不需要进一步的行动
当前解变为不可行	用对偶单纯形恢复可行性
当前解变为非最优	用原始单纯形恢复最优性
当前解变为既非最优也不可行	用广义单纯形法获得新的解

本节研究前 3 种情况. 第 4 种情况是情况 2 和情况 3 的组合, 将在习题 4.5A 的第 6 题来处理.

用例 4.3-2 的 TOYCO 模型来解释不同的方法. TOYCO 模型解决 3 种类型玩具 (玩具火车、玩具卡车和玩具汽车) 的装配问题, 在该装配过程中涉及 3 种操作. 我们希望确定每种玩具的单位数量, 使得收入极大化. 为了方便起见, 这里重新给出该模型及其对偶.

TOYCO 原始问题	TOYCO 对偶问题
max $z = 3x_1 + 2x_2 + 5x_3$	min $w = 430y_1 + 460y_2 + 420y_3$
s.t. $x_1 + 2x_2 + x_3 \leqslant 430$ (操作 1)	s.t. $y_1 + 3y_2 + y_3 \geqslant 3$
$3x_1 + 2x_3 \leqslant 460$ (操作 2)	$2y_1 + 4y_3 \geqslant 2$
$x_1 + 4x_2 \leqslant 420$ (操作 3)	$y_1 + 2y_2 \geqslant 5$
$x_1,\ x_2,\ x_3 \geqslant 0$	$y_1,\ y_2,\ y_3 \geqslant 0$
最优解:	最优解:
$x_1 = 0, x_2 = 100, x_3 = 230, z = \$1\ 350$	$y_1 = 1, x_2 = 2, y_3 = 0, w = \$1\ 350$

原始问题的相应最优单纯形表如下:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	4	0	0	1	2	0	1 350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_6	2	0	0	-2	1	1	20

4.5.1 影响可行性的变化

当前最优解的可行性可能会受到的影响, 仅当 (1) 约束右端项发生变化, 或 (2) 模型中加入一个新的约束. 在这两种情况下, 当最优表中的右端项至少一个元素变为负值, 就会出现不可行性, 也就是说, 当前基变量一个或多个变为负值.

右端项发生变化 这种变化需要用 4.2.4 节的公式 1 重新计算表的右端项:

$$\text{第 } i \text{ 次迭代中表的新右端项} = \text{第 } i \text{ 次迭代中的逆矩阵} \times \text{约束新的右端项}$$

请注意表的右端项给出基变量的值.

例 4.5-1

情形 1 假定 TOYCO 公司打算扩展它的装配线以增加操作 1、操作 2 和操作 3 的日生产能力 40%, 分别达到 602 分钟、644 分钟和 588 分钟. 这种变化将如何影响总收入?

按这些增量, 最优表中发生的改变仅仅是在约束 (和最优目标值) 的右端项. 因此, 新的基本解的计算如下:

$$\begin{pmatrix} x_2 \\ x_3 \\ x_6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 602 \\ 644 \\ 588 \end{pmatrix} = \begin{pmatrix} 140 \\ 322 \\ 28 \end{pmatrix}$$

因此, 当前基变量 x_2, x_3, x_6 保持可行, 其值分别为 140, 322, 28. 相应的最优收入是 1 890 美元, 它比当前的收入 1 350 美元多出 540 美元.

情形 2 尽管新的解从增加收入的观点来看更吸引人, 但 TOYCO 公司认识到, 它的实施可能需要时间. 因此, 另一项提议是将操作 3 生产能力的松弛变量 ($x_6 = 20$ 分钟) 全部换成操作 1 生产能力的松弛变量. 这种变化将如何影响最优解?

这三种操作的生产能力分别达到 450 分钟、460 分钟和 400 分钟. 计算的结果是

$$\begin{pmatrix} x_2 \\ x_3 \\ x_6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 450 \\ 460 \\ 400 \end{pmatrix} = \begin{pmatrix} 110 \\ 230 \\ -40 \end{pmatrix}$$

这个结果是不可行的, 因为 $x_6 = -40$, 需要用对偶单纯形法恢复可行性. 首先, 我们修正由阴影列所显示表的右端项. 注意, 相应的 z 值是 $z = 3 \times 0 + 2 \times 110 + 5 \times 230 = 1\,370$ 美元.

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	4	0	0	1	2	0	1 370
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	110
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_6	2	0	0	-2	1	1	-40

从这个对偶单纯形来看, x_6 离基, x_4 进基, 它将产生如下最优可行表 (在一般情况下, 对偶单纯形可能需要多次迭代才能恢复可行性).

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	5	0	0	0	$\frac{5}{2}$	$\frac{1}{2}$	1 350
x_2	$\frac{1}{4}$	1	0	0	0	$\frac{1}{4}$	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_4	-1	0	0	1	$-\frac{1}{2}$	$-\frac{1}{2}$	20

最优解 (依照 x_1, x_2, x_3) 保持与原模型相同的结果. 这意味着, 在这种情况下, 所提出的生产能力分配的改变并没有带来好处, 因为它仅将操作 3 的剩余生产能力全部变为操作 1 的剩余生产能力. 这个结论说明, 操作 2 是瓶颈, 将剩余生产能力全部用于操作 2, 可能会带来好处 (见习题 4.5A 的第 1 题). 选择操作 2 优于操作 1 也从这样一个事实得到支持, 操作 2 的对偶价格 (2 美元/分钟) 高于操作 1(= 1 美元/分钟).

习题 4.5A

1. 在 4.5 节开始给出的 TOYCO 模型中, 将操作 3 的 20 分钟剩余生产能力分配给操作 2(而不是原题中的操作 1), 能带来更大的好处吗?

2. 假定 TOYCO 打算按照下面的情况改变 3 种操作的生产能力：

$$\textbf{(a)}\begin{pmatrix}460\\500\\400\end{pmatrix}\quad\textbf{(b)}\begin{pmatrix}500\\400\\600\end{pmatrix}\quad\textbf{(c)}\begin{pmatrix}300\\800\\200\end{pmatrix}\quad\textbf{(d)}\begin{pmatrix}450\\700\\350\end{pmatrix}$$

用后最优分析确定每种情况的最优解.

3. 考虑例 2.1-1 中的 Reddy Mikks 模型. 它的最优表由例 3.3-1 中给出. 如果原料 M1 和 M2 的日可用量分别增加到 28 吨和 8 吨, 用后最优分析确定新的最优解.

***4.** Ozark 农场有 20 000 只雏鸡, 在市场出售前喂它到 8 周大. 每只雏鸡的周饲料用量依照如下时间表变化：

周	1	2	3	4	5	6	7	8
磅/只	0.26	0.48	0.75	1.00	1.30	1.60	1.90	2.10

为了使雏鸡在 8 周达到预想的重量, 饲料必须满足特殊的营养需求. 虽然一般说来饲料有很多种, 但为简化起见, 我们仅限制模型在 3 种饲料上：石灰石、玉米和大豆粉. 营养需要也限于 3 种：钙、蛋白质和纤维. 下表概括了所选择原料的营养含量以及相应成本的数据.

原　料	每磅的营养量 (磅)			
	钙	蛋白质	纤维	成本 (美元/磅)
石灰石	0.380	0.00	0.00	0.12
玉　米	0.001	0.09	0.02	0.45
大豆粉	0.002	0.50	0.08	1.60

饲料配制必须包括

(a) 至少 0.8% 但不超过 1.2% 的钙;

(b) 至少 22% 的蛋白质;

(c) 至多 5% 的粗纤维.

对于 1 周饲料用线性规划求解, 然后用后最优分析制定其余 7 周饲料的最优安排表.

5. 证明：习题 3.6C(第 3 章) 第 12 题的 100% 可行性规则可由如下条件导出：

$$\begin{pmatrix}\text{最优}\\\text{逆矩阵}\end{pmatrix}\begin{pmatrix}\text{原始右端}\\\text{项向量}\end{pmatrix}\geqslant 0$$

6. *既影响最优性又影响可行性情况的后最优分析*. 假定在 Reddy Mikks 模型中, 你知道下面的变化同时发生：每吨外、内墙涂料的收入分别是 1 000 美元和 4 000 美元, 原料 M1 和 M2 每天的最大可用量分别是 28 吨和 8 吨.

(a) 证明：所提出的变化将导致当前最优解既不最优也不可行.

(b) 用*广义单纯形算法*(见 4.4.2 节) 求新的最优可行解.

增加新约束　对于一个已有的模型, 增加新约束可能导致两种情况之一发生：

(1) 新约束是多余的, 这意味着当前的最优解满足新约束, 因此, 可以从模型中完全去掉.

(2) 当前解不满足新的约束, 在这种情况下, 使用对偶单纯形法恢复可行性.

注意：增加一个新约束并不一定能够改善当前最优目标值.

例 4.5-2

情形 1 假定 TOYCO 正在改变它的玩具设计, 其改变将需要在装配线上增加第 4 种操作. 新操作每天的生产能力是 500 分钟, 并且对于 3 种产品的单位操作时间分别是 3 分钟、1 分钟和 1 分钟. 研究新操作对于最优解的影响.

操作 4 的约束是

$$3x_1 + x_2 + x_3 \leqslant 500$$

这个约束是多余的, 因为当前最优解 $x_1 = 0, x_2 = 100, x_3 = 230$ 满足此约束. 因此, 当前最优解保持不变.

情形 2 假定 TOYCO 将第 4 种操作的单位时间分别改变为 3 分钟、3 分钟和 1 分钟. 模型中所有其余的数据保持不变, 最优解将如何改变?

操作 4 的约束是

$$3x_1 + 3x_2 + x_3 \leqslant 500$$

当前的最优解不满足上述约束, 因此, 新的约束必须增加到下面的当前最优表中 (x_7 是松弛变量):

基	x_1	x_2	x_3	x_4	x_5	x_6	x_7	解
z	4	0	0	1	2	0	0	1 350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	0	230
x_6	2	0	0	-2	1	1	0	20
x_7	3	3	1	0	0	0	1	500

此表表明, $x_7 = 500$, 它并不与表中其他元素 x_2 和 x_3 的值相一致. 其原因是基变量 x_2 和 x_3 还没有在新的约束中被替换出去. 这种替换由下面的运算完成:

$$\text{新的 } x_7 \text{ 行} = \text{旧的 } x_7 \text{ 行} - \{3 \times (x_2 \text{ 行}) + 1 \times (x_3 \text{ 行})\}$$

这个运算在新的约束中与下面的替换完全相同.

$$x_2 = 100 - \left(-\tfrac{1}{4}x_1 + \tfrac{1}{2}x_4 - \tfrac{1}{4}x_5\right)$$
$$x_3 = 230 - \left(\tfrac{3}{2}x_1 + \tfrac{1}{2}x_5\right)$$

因此, 新表给出如下:

基	x_1	x_2	x_3	x_4	x_5	x_6	x_7	解
z	4	0	0	1	2	0	0	1 350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	0	230
x_6	2	0	0	-2	1	1	0	20
x_7	$\frac{9}{4}$	0	0	$-\frac{3}{2}$	$\frac{1}{4}$	0	1	-30

应用对偶单纯形法将产生新的最优解 $x_1 = 0$, $x_2 = 90$, $x_3 = 230$, $z = 1\ 330$(请验证!). 这个解表明, 增加操作 4 后收入从 1 350 美元降低到 1 330 美元.

习题 4.5B

1. 在 TOYCO 模型中, 假定第 4 种操作有如下规范: 一天接 480 分钟工作量, 其最大生产率或者是 120 个单位的产品 1, 或者是 480 个单位的产品 2, 或者是 240 个单位的产品 3. 确定最优解, 假定一天的生产能力限制在

*(a) 570 分钟. (b) 548 分钟.

2. *次要约束*. 不用全部约束求解问题, 我们可以通过找出所谓的*次要约束*来开始求解. 这些约束是我们所怀疑的在最优解上有最少限制的约束. 用其余 (主要) 的约束求解模型. 然后, 我们一次增加一个次要约束. 如果次要约束满足已有的最优解, 则将它放弃. 重复这个过程直到所有的次要约束都计算完.

用上述方法求解下面的线性规划:

$$
\begin{aligned}
\max \quad & z = 5x_1 + 6x_2 + 3x_3 \\
\text{s.t.} \quad & 5x_1 + 5x_2 + 3x_3 \leqslant 50 \\
& x_1 + x_2 - x_3 \leqslant 20 \\
& 7x_1 + 6x_2 - 9x_3 \leqslant 30 \\
& 5x_1 + 5x_2 + 5x_3 \leqslant 35 \\
& 12x_1 + 6x_2 \leqslant 90 \\
& x_2 - 9x_3 \leqslant 20 \\
& x_1, x_2, x_3 \geqslant 0
\end{aligned}
$$

4.5.2 影响最优性的变化

这一节考虑可能影响当前解最优性的两种特殊情形:

(1) 原目标系数发生了变化; (2) 在模型中增加了新的经济活动 (变量).

目标函数系数的变化 这些变化只影响解的最优性. 因此, 此类变化需要依照下面的过程重新计算 z 行系数 (简约费用):

(1) 用 4.2.3 节中的方法 2 计算对偶值.

(2) 用 4.2.4 节公式 2 的新对偶值确定新的简约费用 (z 行系数).

两种情况将导致:

(1) 新的 z 行满足最优性条件. 解保持不变 (然而, 最优目标值可能变化).

(2) 最优条件不满足. 应用 (原始) 单纯形法恢复最优性.

例 4.5-3

情形 1 在 TOYCO 模型中, 假定公司有新的定价策略来适应竞争需要. 在新的策略下, 玩具火车、玩具卡车和玩具汽车的单位收入分别是 2 美元、3 美元和 4 美元. 最优解会受到什么样的影响?

新的目标函数是

$$\max\ z = 2x_1 + 3x_2 + 4x_3$$

因此,

$$\text{基变量 } x_2,\ x_3,\ x_6 \text{ 新的目标系数} = (3, 4, 0)$$

用 4.2.3 节中的方法 2, 计算对偶值如下:

$$(y_1, y_2, y_3) = (3, 4, 0)\begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix} = \left(\tfrac{3}{2}, \tfrac{5}{4}, 0\right)$$

z 行系数由对偶约束左端项与右端项的差来确定 (4.2.4 节, 公式 2). 不必重新计算基变量 x_2, x_3, x_6 的目标行的系数, 因为无论目标系数如何变化, 它们总是等于零 (请验证!).

$$x_1 \text{ 的简约费用} = y_1 + 3y_2 + y_3 - \mathbf{2} = \tfrac{3}{2} + 3\left(\tfrac{5}{4}\right) + 0 - \mathbf{2} = \tfrac{13}{4}$$
$$x_4 \text{ 的简约费用} = y_1 - 0 = \tfrac{3}{2}$$
$$x_5 \text{ 的简约费用} = y_2 - 0 = \tfrac{5}{4}$$

注意: 在所修正的目标函数新的系数中, 第一个对偶约束的右端项是 2.

计算表明, 当前解中, $x_1 = 0$ 辆玩具火车, $x_2 = 100$ 辆玩具卡车, $x_3 = 230$ 辆玩具汽车, 保持最优. 相应的新收入计算为 $2 \times 0 + 3 \times 100 + 4 \times 230 = 1\,220$ 美元. 新的定价策略没有任何优点, 因为它导致了更低的收入.

情形 2 现在假定 TOYCO 目标函数变成

$$\max\ z = 6x_1 + 3x_2 + 4x_3$$

最优解将如何变化呢?

我们有

$$(y_1, y_2, y_3) = (3, 4, 0)\begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix} = \left(\tfrac{3}{2}, \tfrac{5}{4}, 0\right)$$

$$x_1 \text{ 的简约费用} = y_1 + 3y_2 + y_3 - \mathbf{6} = \frac{3}{2} + 3\left(\frac{5}{4}\right) + 0 - \mathbf{6} = -\frac{3}{4}$$

$$x_4 \text{ 的简约费用} = y_1 - 0 = \frac{3}{2}$$

$$x_5 \text{ 的简约费用} = y_2 - 0 = \frac{5}{4}$$

x_1 新的简约费用表明, 当前解不是最优的.

为了确定新的解, z 行变化如下, 由下表中突出显示部分所示:

基	x_1	x_2	x_3	x_4	x_5	x_6	解
z	$-\frac{3}{4}$	0	0	$\frac{3}{2}$	$\frac{5}{4}$	0	1 220
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
x_6	2	0	0	-2	1	1	20

在阴影单元中显示的元素是非基变量 x_1, x_4, x_5 新的简约费用. 所有其余的元素与原始最优表相同. 然后, 新的最优解是由 x_1 进基和 x_6 离基确定, 它产生 $x_1 = 10$, $x_2 = 102.5$, $x_3 = 215$, 并且 $z = 1\ 227.50$(请验证!). 尽管新的解推荐生产所有的 3 种玩具, 但最优收入还是小于只制造 2 种玩具时收入.

习题 4.5C

1. 对于下面各个目标函数, 研究 TOYCO 解的最优性. 如果解发生变化, 用后最优分析确定新的最优值. (4.5 节初给出了 TOYCO 的最优表.)

(a) $z = 2x_1 + x_2 + 4x_3$ **(b)** $z = 3x_1 + 6x_2 + x_3$ **(c)** $z = 8x_1 + 3x_2 + 9x_3$

2. 对于下面各个目标函数, 研究 Reddy Mikks 解 (例 4.3-1) 的最优性. 如果解发生变化, 用后最优分析确定新的最优值. (例 3.3-1 中给出了模型的最优表.)

*__(a)__ $z = 3x_1 + 2x_2$ **(b)** $z = 8x_1 + 10x_2$ *__(c)__ $z = 2x_1 + 5x_2$

3. 证明: 100% 最优性规则 (见第 3 章习题 3.6D 的第 8 题) 可由极大化问题的 (简约费用)$\geqslant 0$ 以及极小化问题的 (简约费用)$\leqslant 0$ 推导出.

增加新活动 在线性规划模型中增加一项新活动等价于增加一个新变量. 直观地说, 一项新活动是值得增加的, 只有当它有利可图, 也就是说, 如果它改善目标函数的最优值. 可以使用 4.2.4 节的公式 2 计算新变量的简约费用来检验这个条件. 如果新的活动满足最优性条件, 则此项活动无利可图; 否则, 承担新活动是有利的.

例 4.5-4

TOYCO 认识到, 玩具火车并不在当前的产品中, 因为它们无利可图. 公司打算用新的产品玩具消防车替代玩具火车, 并且玩具消防车可在已有的设备上装配. TOYCO 估计每辆玩具消防车的收入可达到 4 美元, 而且每单位的装配时间在操作 1 和操作 2 上都是 1 分钟, 在操作 3 上是 2 分钟. 这种改变将如何影响解?

令 x_7 表示新的消防车产品. 已知 $(y_1, y_2, y_3) = (1, 2, 0)$ 是最优对偶值, 我们得到

$$x_7 \text{ 的简约费用} = 1y_1 + 1y_2 + 2y_3 - 4 = 1 \times 1 + 1 \times 2 + 2 \times 0 - 4 = -1$$

这个结果表明, 在最优基本解中包含 x_7 是有利可图的. 为了获得新的最优解, 我们首先用 4.2.4 节的公式 1 计算它的列约束如下:

$$x_7 \text{ 的约束列} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

因此, 当前的单纯形表必须修正, 其结果如下:

基	x_1	x_2	x_3	x_7	x_4	x_5	x_6	解
z	4	0	0	-1	1	2	0	1 350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{4}$	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	230
x_6	2	0	0	1	-2	1	1	20

新的最优解可以通过令 x_7 进入基本解确定, 在这种情况下 x_6 必须离基. 新的解是 $x_1 = 0$, $x_2 = 0$, $x_3 = 125$, $x_7 = 210$, 且 $z = 1\ 465$ 美元 (请验证!), 它增加收入 115 美元.

习题 4.5D

*1. 在原始 TOYCO 模型中, 玩具火车不是最优生产规划的部分. 公司认识到由于市场竞争激烈, 无法提高玩具的单位价格. 相反, 公司希望集中改进自身装配操作. 这样就必须减少每种玩具在每种操作上的装配时间, 减少的时间为指定的百分比 $p\%$. 确定 p 值, 使得玩具火车恰好有利可图. (TOYCO 模型的最优表在 4.5 节初给出.)

2. 在 TOYCO 模型中, 假定公司能够减少玩具火车在操作 1、操作 2 和操作 3 上的单位时间, 分别从当前的 1 分钟、3 分钟和 1 分钟减少到 0.5 分钟、1 分钟和 0.5 分钟. 玩具火车的单位收入保持不变, 仍然是 3 美元. 确定新的最优解.

3. 在 TOYCO 模型中, 假定新玩具 (消防车) 在操作 1、操作 2 和操作 3 上分别需要 3 分钟、2 分钟和 4 分钟. 每单位的收入如下给出, 确定最优解.

*(a) 5 美元. **(b)** 10 美元.

4. 在 Reddy Mikks 模型中, 公司正在考虑生产廉价牌子的外墙涂料, 每生产一吨这样的涂料需要原料 M1 和 M2 每种 0.75 吨. 市场条件仍然表明, 内墙涂料的日需求量不超过两种类型外墙涂料的需求量加上 1 吨. 新的外墙涂料每吨的收入是 3 500 美元. 确定新的最优解. (模型在例 4.5-1 中得到了解释, 而且它的最优表由例 3.3-1 给出.)

参 考 文 献

Bradley, S., A. Hax, and T. Magnanti, *Applied Mathematical Programming*, Addison-Wesley, Reading, MA, 1977.

Bazaraa, M., J. Jarvis, and H. Sherali, *Linear Programming and Network Flows*, 2nd ed., Wiley, New York, 1990.

Diwckar, U., *Introduction to Applied Optimization*, Kluwer Academic Publishers, Boston, 2003.

Nering, E., and A.Tucker, *Linear Programming and Related Problems*, Academic Press, Boston, 1992.

Vanderbei, R., *Linear Programming: Foundation and Optimization*, 2nd ed., Kluwer Academic Publishers, Boston MA, 2001.

第 5 章　各种运输模型

本章导读　运输模型是一类特殊的线性规划, 研究如何把商品从*起点*(如工厂)运送到*终点*(如商场), 目标是确定一项运输计划, 在满足供需约束的条件下使得总的运输费用最小. 运输模型的应用可以扩展到库存控制、招聘计划、人员指派等领域.

学习本章需要注意的是, 运输算法的步骤与单纯形方法的步骤完全一致. 另外一点是运输算法发展于运筹学早期, 主要是提高手工计算效率. 现在, 随着计算机能力的增强, 不再强调使用这些简便方法. 实际上, 本章的方法从严格意义下不会出现在商业软件上. 然而, 本章介绍的内容表明, 专用的运输模型表上作业法对一些简单问题的建模特别有用 (和我们常用的带有明确目标函数和约束的线性规划模型完全不同). 尤其是, 采用 Excel 规划求解进行运输表上计算可以简化问题的求解. 本章还介绍了一些很有意思的思想, 说明如何运用线性规划的基本理论来得到简便的计算方法.

你会发现, TORA 的教学模块对于理解运输问题的详细算法很有帮助, 这些程序模块能够让你通过即时的反馈来对计算的合理性做出判断.

本章包括 1 个实际应用、12 个例题、1 个 Excel 规划求解模型、4 个 AMPL 模型、46 个节后习题和 5 个案例. 这些案例都收录在本书 (下册) 的附录 E 中. 本章有关 AMPL/Excel/Solver/TORA 的程序放在下载资料的文件夹 ch5Files 中.

实际应用——澳大利亚旅游委员会关于旅游交易会的会面安排问题

澳大利亚旅游委员会 (Australian Tourist Commission, ATC) 在全球各地组织旅游产品交易会, 为澳大利亚的旅游销售商提供一种平台, 向国际购买商推销旅店、旅游线路、交通等旅游产品. 交易会期间, 销售商设立展台, 按照事先安排好的会面计划接待购买商. 因为每次交易会的时间有限, 而且销售商和购买商的数量非常大 (1997 年在墨尔本举办的交易会上, 吸引了 620 个销售商和 700 个购买商), 因此 ATC 希望在交易会之前精心安排好买卖双方的会面, 使得最大程度地符合双方的兴趣. 本问题的模型让供需双方都得到了最大的满足. 第 24 章的案例 3 给出了本问题的详细案例分析.

5.1　运输模型的定义

一般的运输问题可用图 5.1 的网络图表示. 图中有 m 个起点和 n 个终点, 每个用**节点** (node) 表示. 连接起点和终点的路线用**弧** (arc) 表示. 连接起点 i 到终点 j 的弧 (i,j) 带有两个信息：每单位运输费用为 c_{ij}, 运输量为 x_{ij}. 起点 i 的供应量 a_i, 终点 j 的需求量为 b_j. 这一模型的目标是确定未知变量 x_{ij}, 在满足供应和需求约束的情况下, 使得运输总费用最小.

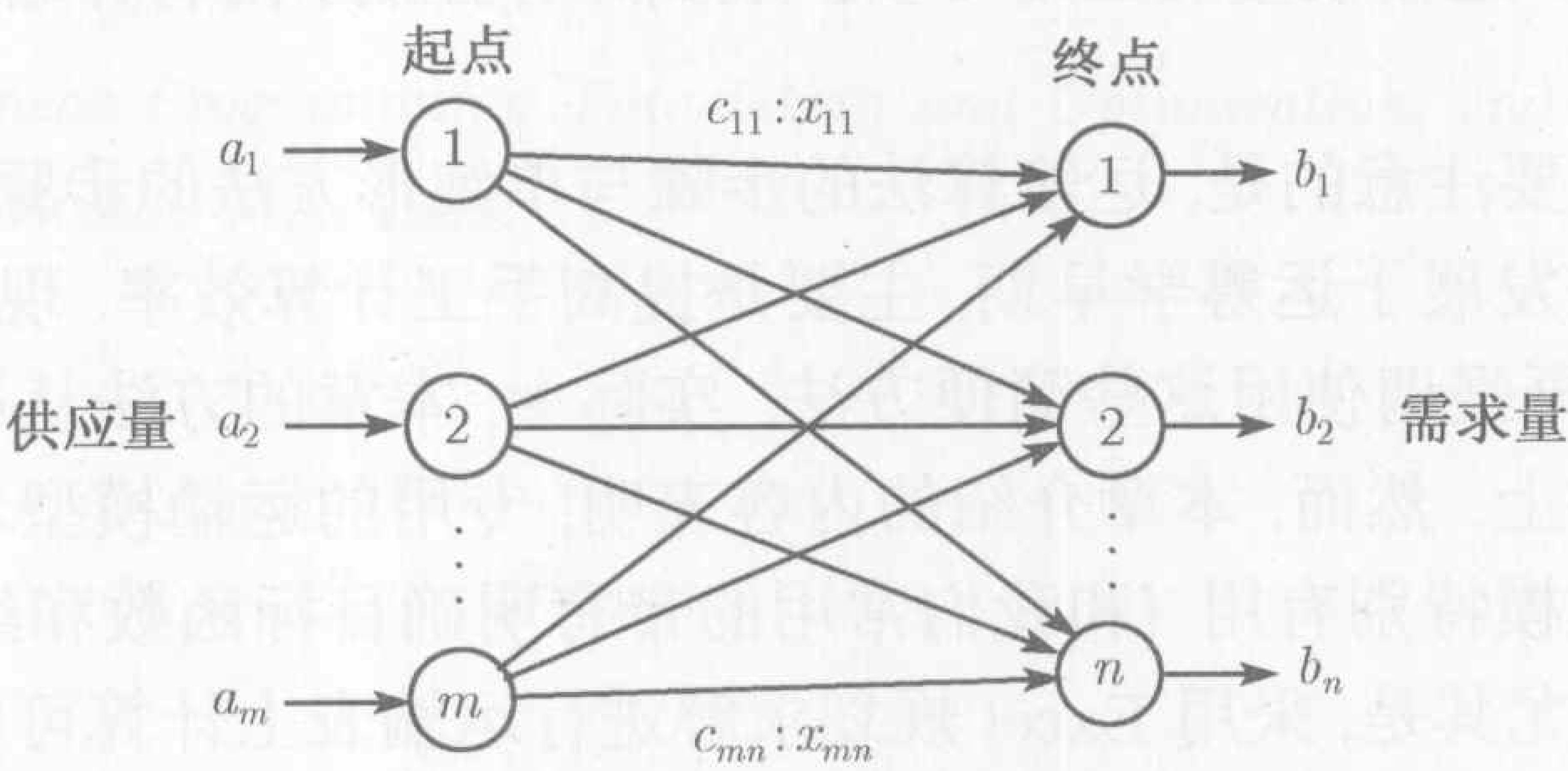


图 5.1　用节点和弧表示的运输模型

例 5.1-1

MG 汽车制造公司有 3 个生产厂, 分别位于洛杉矶、底特律和新奥尔良, 在丹佛和迈阿密设有 2 个主要分销中心. 这 3 个厂下个季度的生产能力分别为 1 000, 1 500, 1 200 辆汽车. 2 个分销中心的季度需求分别是 2 300 和 1 400 辆汽车. 生产厂与分销中心之间的距离里程如表 5.1 所示.

表 5.1　距离里程表

	丹佛	迈阿密
洛杉矶	1 000	2 690
底特律	1 250	1 350
新奥尔良	1 275	850

负责运送这些汽车的卡车运输公司按照每辆汽车每英里 8 美分收取运费, 每辆汽车不同路线上的运费, 四舍五入换算成美元数, 如表 5.2 所示.

表 5.2　每辆汽车的运输费用

	丹佛 (1)	迈阿密 (2)
洛杉矶 (1)	$80	$215
底特律 (2)	$100	$108
新奥尔良 (3)	$102	$68

本问题的线性规划模型为

$$
\begin{aligned}
\min \quad & z = 80x_{11} + 215x_{12} + 100x_{21} + 108x_{22} + 102x_{31} + 68x_{32} \\
\text{s.t.}\quad & x_{11} + x_{12} = 1\,000 \quad (\text{洛杉矶}) \\
& x_{21} + x_{22} = 1\,500 \quad (\text{底特律}) \\
& x_{31} + x_{32} = 1\,200 \quad (\text{新奥尔良}) \\
& x_{11} + x_{21} + x_{31} = 2\,300 \quad (\text{丹佛}) \\
& x_{12} + x_{22} + x_{32} = 1\,400 \quad (\text{迈阿密}) \\
& x_{ij} \geqslant 0, i = 1,2,3, \quad j = 1,2
\end{aligned}
$$

因为从 3 个起点的总供应量 (=1 000+1 500+1 200=3 700 辆汽车) 等于 2 个分销中心的总需求 (=2 300+1 400 辆汽车), 因此这些约束条件都是等式.

这个线性规划模型可以运用单纯形方法求解. 然而由于运输问题的特殊结构, 采用**运输表** (transportation tableau) 的方法来求解此问题更为方便, 如表 5.3 所示.

表 5.3 MG 汽车运输模型

	丹佛	迈阿密	供应量
洛杉矶	80 x_{11}	215 x_{12}	**1 000**
底特律	100 x_{21}	108 x_{22}	**1 500**
新奥尔良	102 x_{31}	68 x_{32}	**1 200**
需求量	**2 300**	**1 400**	

图 5.2 中, 最优解 (用 TORA[①]得到) 为从洛杉矶运送 1 000 辆汽车到丹佛, 从底特律运送 1 300 辆汽车到丹佛; 从底特律运送 200 辆汽车到迈阿密, 从新奥尔良运送 1 200 辆汽车到迈阿密. 相应的最小运输费用为: $1\,000 \times \$80 + 1\,300 \times \$100 + 200 \times \$108 + 1\,200 \times \$68 = \$313\,200$.

运输模型的平衡 运输算法假定模型是平衡的, 即总需求量等于总供应量. 如果模型不是平衡的, 我们总可以通过增加一个虚设起点或一个虚设终点以使得模型达到平衡.

---

① 为了运用 TORA, 从 Main Menu (主菜单) 中选择 Transportation Model (运输模型). 从 SOLVE/MODIFY 菜单选择 Solve ⇒ Final Solution 得到最优解. 运输模型迭代解的细节见 5.3.3 节.

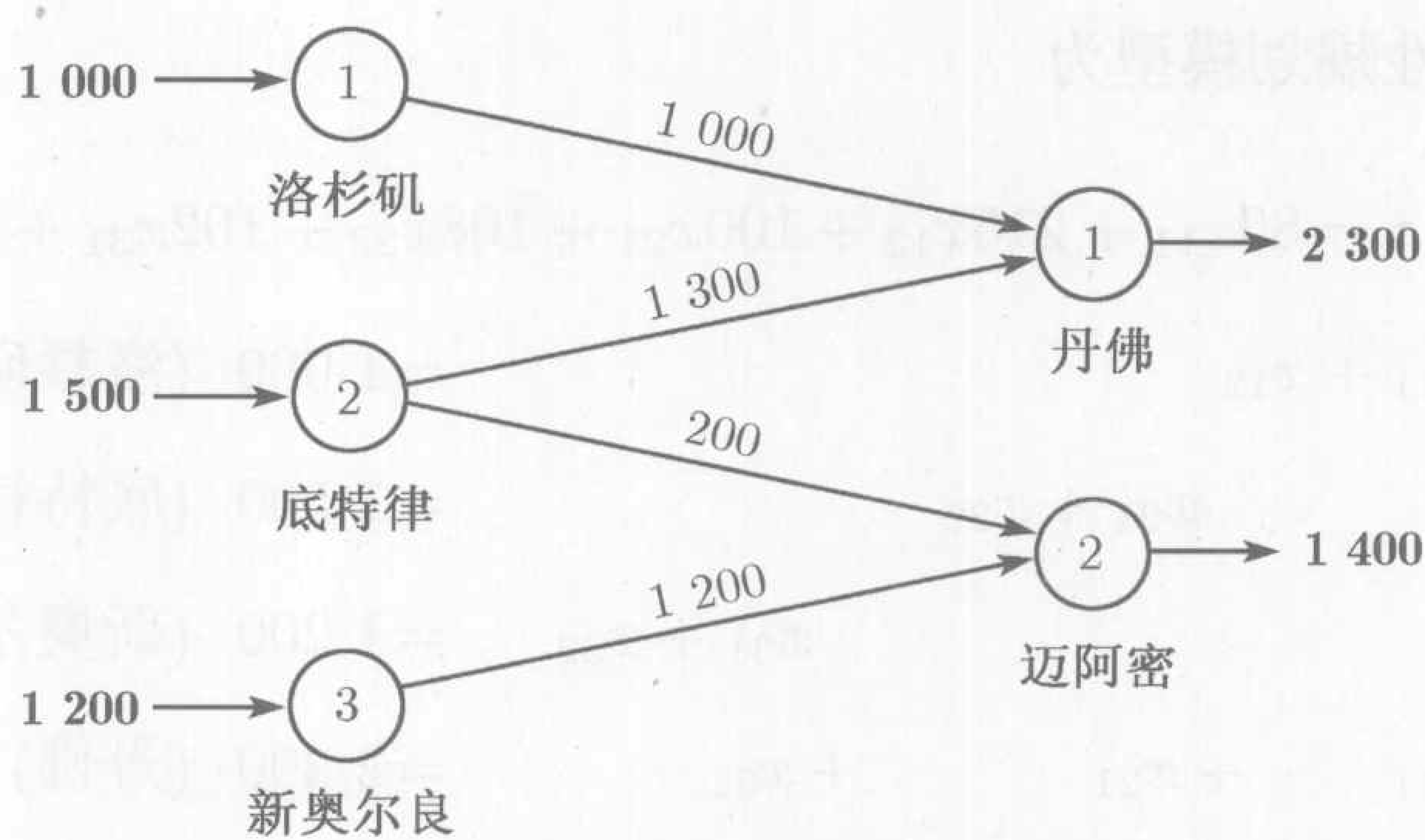


图 5.2 MG 汽车运输模型的最优解

例 5.1-2

在 MG 汽车运输模型中, 假设底特律工厂的生产能力为 1 300 辆汽车 (而不是 1 500 辆汽车). 总供应量 (=3 500 辆汽车) 小于总需求 (=3 700 辆汽车), 这意味着丹佛和迈阿密的部分需求将得不到满足.

由于需求量大于供应量, 我们增加一个生产能力为 200 辆汽车 (= 3 700−3 500) 的虚设起点 (工厂) 来平衡运输模型. 因为这样的工厂并不存在, 单位运输费用从虚设起点到终点的运输费为 0.

表 5.4 给出了平衡后的模型及其最优解. 这个最优解表明虚设工厂运输了 200 辆汽车到迈阿密, 即迈阿密的 1 400 辆汽车的需求中将缺少 200 辆.

表 5.4 带有虚设工厂的MG汽车运输模型

	丹佛	迈阿密	供应量
洛杉矶	80 **1 000**	215	**1 000**
底特律	100 **1 300**	108	**1 300**
新奥尔良	102	68 **1 200**	**1 200**
虚设工厂	0	0 **200**	**200**
需求量	**2 300**	**1 400**	

可以肯定的是, 一个具体的终点不会因为从虚设起点到这个终点的运输费用高而出现需求短缺. 例如, 把从虚设起点到迈阿密的运输费用定为$1 000 仍然可以使迈阿密的需求得到满足. 当然, 不能对所有的终点都使用这种“技巧”, 因为在这个系统中总有某些终点得不到满足.

为了说明供应量大于需求量的情况, 我们假设丹佛的需求只有 1 900 辆汽车. 在这种情况下, 我们需要增加一个虚设分销中心来“接收”多余的汽车, 同样地, 从工厂到该虚设终点的运输费用为 0, 除非我们指定某个工厂完全“运出去”, 在这种情况下, 我们必须把从这个指定工厂到虚设分销中心的单位运输费用设置成充分大.

表 5.5 给出了一个新的模型以及 (用 TORA 计算的) 最优解. 这个最优解表明, 底特律工厂剩余了 400 辆汽车.

表 5.5 带有虚设终点的 MG 汽车运输模型

	丹佛	迈阿密	虚设终点	
洛杉矶	80 **1 000**	215	0	**1 000**
底特律	100 **900**	108 **200**	0 **400**	**1 500**
新奥尔良	102	68 **1 200**	0	**1 200**
需求量	**1 900**	**1 400**	**400**	

习题 5.1A[①]

1. 判断正误:

(a) 为了平衡运输模型, 必须同时增加一个虚设起点和一个虚设终点.

(b) 运往虚设终点的运输量代表运输起点的剩余量.

(c) 从虚设起点运出的货物量表示接受终点的短缺量.

2. 在下列情况下, 决定是否需要增加虚设起点或虚设终点来平衡模型.

(a) 供应量: $a_1 = 10, a_2 = 5; a_3 = 4, a_4 = 6$

需求量: $b_1 = 10, b_2 = 5, b_3 = 7, b_4 = 9$

(b) 供应量: $a_1 = 30, a_2 = 44$

需求量: $b_1 = 25, b_2 = 30, b_3 = 10$

3. 在例 5.1-2 中的表 5.4 中, 虚设工厂需要增加在哪里? 当虚设工厂“运输”150 辆汽车到丹佛和 50 辆汽车到迈阿密时, 这个解意味着什么?

***4.** 在例 5.1-2 中的表 5.5 中增加一个虚设终点, 假定底特律工厂必须运出它的所有产品, 如何在模型中体现这个限制条件?

5. 在例 5.1-2 中, 假如出现了需求大于供应的情况 (表 5.4), 对没有送到丹佛和底特律的每辆汽车分别罚款$200 和$300, 另外, 从洛杉矶工厂不能运送汽车到迈阿密分销中心. 建立这一模型, 并求出本问题的最优运输方案.

① 求解本习题时, 你可以用 TORA 程序求最优解. 运输问题的 AMPL 和 Excel 规划求解程序将在 5.3.2 节介绍.

***6.** 有 3 个发电厂向 3 个城市供电, 其发电量分别为 25、40 和 30 兆千瓦时. 3 个城市的最大需求量估计为 30、35 和 25 兆千瓦时. 这 3 个城市的每兆千瓦时电价如表 5.6 所示.

表 5.6 第 6 题的电价表(元/兆千瓦时)

		城市		
		1	2	3
电厂	1	$600	$700	$400
	2	$320	$300	$350
	3	$500	$480	$450

在 8 月份期间, 这 3 个城市的每一个都会增加 20%的用电需求量, 这部分电量可以从别的电网以每兆千瓦时$1 000 的额外费用购得. 这个电网不与第 3 个城市相连, 电力公司希望确定增加电量的最经济的销售和采购计划.

(a) 把此问题表达为一个运输模型. **(b)** 为电力公司制定一个最优销售计划.

(c) 确定 3 个城市每个的额外购电费用.

7. 求解第 6 题, 并假设电网有 10%的输电损耗.

8. 3 个炼油厂日炼油能力分别为 600 万、500 万和 800 万加仑, 所供应 3 个分销区域的日需求量分别为 400 万、800 万和 700 万加仑. 汽油通过输油管线被运送到 3 个分销区域. 管线每英里运送 1 000 加仑的运费为 10 美分. 表 5.7 给出了炼油厂到分销区域的里程表. 其中, 炼油厂 1 与分销区域 3 不相连.

(a) 建立相应的运输模型. **(b)** 确定本运输网络的最优运输策略.

表 5.7 第 8 题的里程表

		分销区域		
		1	2	3
炼油厂	1	120	180	—
	2	300	100	80
	3	200	250	120

***9.** 在第 8 题中, 假设炼油厂 3 的日生产能力仅为 600 万加仑, 并且分销区域 1 的需求量必须满足, 分销区域 2 和分销区域 3 每缺少一加仑油有 5 美分的惩罚费用.

(a) 建立相应的运输模型. **(b)** 求出最优运输策略.

10. 在第 8 题中, 假设分销区域 3 的日需求量降低到 400 万加仑, 炼油厂 1 和炼油厂 2 的剩余产品用卡车运输到其他分销区域. 从炼油厂 1 出发的运输费用为每 100 加仑$ 1.50, 从炼油厂 2 出发的运输费用为每 100 加仑$ 2.20. 炼油厂 3 在本厂内可以通过别的化学工艺转移剩余产品.

(a) 建立相应的运输模型. **(b)** 确定最优运输策略.

11. 有 3 个果园向 4 个零售商供应甜橙, 4 个零售商的日需求量分别是 150、150、400 和 100 筐. 这 3 个果园的供应量受正常工作量的限制, 每天的供应能力分别约为 150、200 和 250

筐. 然而, 果园 1 和果园 2 表示如果需要, 可以通过加班来增加供应量, 但果园 3 不采用这种方式. 从果园到零售店每筐甜橙的运输费用如表 5.8 所示.

表 5.8 第 11 题中每筐甜橙的运费表

		零售商			
		1	2	3	4
果园	1	\$1	\$2	\$3	\$2
	2	\$2	\$4	\$1	\$2
	3	\$1	\$3	\$5	\$3

(a) 为该问题建立相应的运输模型. (b) 求解这个问题.

(c) 利用加班工作量, 果园 1 和果园 2 应供应多少筐?

12. 从 3 个分销中心向 5 个经销商运输汽车. 运输费用依据起点到终点的距离而定, 与运输卡车是否满载无关. 如表 5.9 所示, 表中总结了分销中心到经销商之间的里程以及相应的月供应量与需求量 (数字代表汽车数). 运输卡车满载量为 18 辆汽车, 每辆卡车每英里的运输费用为\$25.

表 5.9 第 12 题的里程表及其供需量

		经销商					
		1	2	3	4	5	供应量
分销中心	1	100	150	200	140	35	**400**
	2	50	70	60	65	80	**200**
	3	40	90	100	150	130	**150**
需求量		**100**	**200**	**150**	**160**	**140**	

(a) 建立相应的运输模型. (b) 求最优运输安排方案.

13. 例 5.1-1 中的 MG 汽车制造公司生产 M1、M2、M3 和 M4 四种型号的汽车: 底特律厂生产 M1、M2 和 M4 型, 新奥尔良厂生产 M1 和 M2 型, 洛杉矶厂生产 M3 和 M4 型. 各厂的生产能力与各分销中心的需求量如表 5.10 所示.

表 5.10 第 13 题的运输能力和需求

	车型				总计
	M1	M2	M3	M4	
生产厂					
洛杉矶	—	—	700	300	1 000
底特律	500	600	—	400	1 500
新奥尔良	800	400	—	—	1 200
分销中心					
丹佛	700	500	500	600	2 300
迈阿密	600	500	200	100	1 400

里程表与例 5.1-1 一样, 所有类型汽车的运输费用都是每辆车每英里 8 美分. 另外, 如表 5.11 所示, 可以按照一定的百分比从别的厂满足分销中心对于某些型号汽车的需求量.

表 5.11 第 13 题的可互换模型

分销中心	需求百分比	可互换型号
丹佛	10	M1,M2
	20	M3, M4
迈阿密	10	M1, M3
	5	M2, M4

(a) 建立相应的运输模型. (b) 求出最优运输计划.

(提示: 对应于新的产品组合 [M1, M2]、[M3, M4]、[M1, M3] 和 [M2, M4], 增加 4 个新的终点, 这些新终点的需求量由给定的百分比确定.)

5.2 非传统运输模型

运输模型的应用并不仅限于在不同起点和终点之间运送货物, 本节将介绍运输模型在生产–库存控制和刀具打磨服务方面的两个应用领域.

例 5.2-1 (生产–库存控制)

Boralis 公司生产专业徒步旅行者用的背包. 产品需求通常出现在每年的 3 月至 6 月, Boralis 公司估计这 4 个月的需求量分别为 100, 200, 180, 300 个. 由于公司雇用兼职人员生产背包, 因此每月的生产能力都不一样. 据估计, Boralis 公司从 3 月到 6 月能够生产 50, 180, 280, 270 个. 因为不同月份的生产能力和需求不匹配, 当前月份的需求可能通过以下 3 种方法来满足.

(1) 当月生产.

(2) 以前某个月剩余的产品.

(3) 以后某个月多余的产品 (延期交货).

在第 1 种情况下, 背包的生产费用为每个$40. 第 2 种情况下每个背包每月的储存费用为$0.50. 第 3 种情况下每个背包每月的延期交货惩罚费为$2.00. Boralis 公司希望确定这 4 个月期间的最优生产计划.

通过找出生产–库存问题与运输模型要素之间的下列对应关系, 可以把这个问题建立为一个运输模型:

运输	生产–库存
(1) 起点 i	(1) 生产周期 i
(2) 终点 j	(2) 需求周期 j
(3) 起点 i 的供应量	(3) 生产周期 i 的生产能力
(4) 终点 j 的需求量	(4) 周期 j 的需求量
(5) 从起点 i 到终点 j 的单位运输费用	(5) 周期 i 为周期 j 生产的单位费用 (生产 + 库存 + 惩罚)

表 5.12 给出了所建立的运输模型.

表 5.12 例 5.2-1 的运输模型

	1	2	3	4	生产能力
1	\$40.00	\$40.50	\$41.00	\$41.50	**50**
2	\$42.00	\$40.00	\$40.50	\$41.00	**180**
3	\$44.00	\$42.00	\$40.00	\$40.50	**280**
4	\$46.00	\$44.00	\$42.00	\$40.00	**270**
需求量	**100**	**200**	**180**	**300**	

从周期 i 到周期 j 的单位"运输"费用可计算为:

$$c_{ij}=\begin{cases}\text{周期 } i \text{ 的生产费用}, i=j\\ \text{周期 } i \text{ 的生产费用} + \text{从 } i \text{ 到 } j \text{ 的库存费用}, i<j\\ \text{周期 } i \text{ 的生产费用} + \text{从 } i \text{ 到 } j \text{ 的惩罚费用}, i>j\end{cases}$$

例如,

$$c_{11}=\$40.00$$

$$c_{24}=\$40.00+(\$0.50+\$0.50)=\$41.00$$

$$c_{41}=\$40.00+(\$2.00+\$2.00+\$2.00)=\$46.00$$

最优解如图 5.3 所示. 虚线表示延期交货, 点线表示为后来周期生产产品, 而实线则表示为本周期生产. 总费用为 \$31 455.

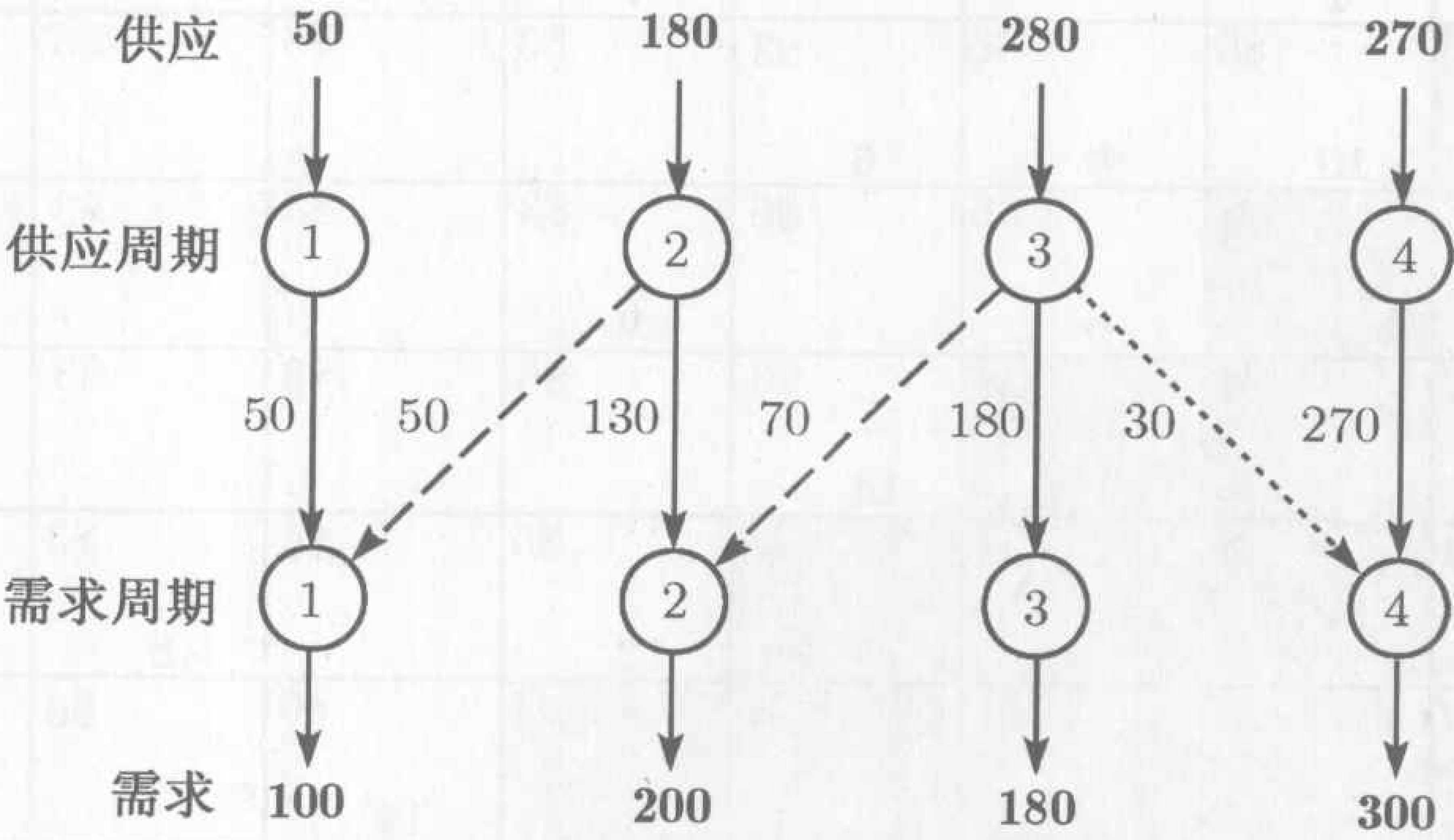


图 5.3 生产–库存模型的最优解

例 5.2-2 (工具打磨)

阿肯色太平洋公司经营一个中等规模的锯木厂. 锯木厂根据每周的安排, 要准备从软松木到硬橡木等不同种类的木材. 根据要锯木材的类型, 一周内每天锯片的需求量也不同, 如下表所示:

日	星期一	星期二	星期三	星期四	星期五	星期六	星期日
(锯片) 需求量	24	12	14	20	18	14	22

工厂可以通过以下方式满足每天的需求:

(1) 以每片 $12 的价格购买新锯片.

(2) 利用一种通宵打磨服务, 磨每片锯片的费用为 $6.

(3) 使用稍慢一点的 2 天打磨服务, 磨每片锯片的费用为 $3.

本问题可以表示成 8 个起点和 7 个终点的运输模型. 终点表示一星期里的 7 天. 模型的起点定义如下: 起点 1 对应于购买新锯片, 在极端情况下, 可以提供足够多的锯片以满足所有 7 天的需求 $(= 24+12+14+20+18+14+22 = 124)$. 起点 2 到 8 对应于一星期中的 7 天. 这些起点的供应量等于相应每天所使用锯片的数量. 例如, 起点 2(即星期一) 可以提供用过的锯片数量等于星期一对锯片的需求量. 本模型的单位“运输费用”, 依据是否购买新锯片、通宵打磨服务或晚 2 天打磨服务分别为$12、$6 或$3. 需要注意的是, 通宵打磨服务意味着第 i 天*结束*时所用过的锯片可以在第 $(i+1)$ 天或第 $(i+2)$ 天的*开始*时使用, 慢的 2 天打磨服务的锯片可以在第 $(i+3)$ 天的*开始*时使用. “处理”列是平衡模型所需要的一个虚设终点. 整个模型及其解如表 5.13 所示.

表 5.13 用运输模型表示工具打磨问题

	1 星期一	2 星期二	3 星期三	4 星期四	5 星期五	6 星期六	7 星期日	8 处理	
1-新锯片	$12 **24**	$12 **2**	$12	$12	$12	$12	$12	$0 **98**	**124**
2-星期一	M	$6 **10**	$6 **8**	$3 **6**	$3	$3	$3	$0	**24**
3-星期二	M	M	$6 **6**	$6	$3 **6**	$3	$3	$0	**12**
4-星期三	M	M	M	$6 **14**	$6	$3	$3	$0	**14**
5-星期四	M	M	M	M	$6 **12**	$6	$3 **8**	$0	**20**
6-星期五	M	M	M	M	M	$6 **14**	$6	$0 **4**	**18**
7-星期六	M	M	M	M	M	M	$6 **14**	$0	**14**
8-星期日	M	M	M	M	M	M	M	$0 **22**	**22**
	24	**12**	**14**	**20**	**18**	**14**	**22**	**124**	

本问题还存在另外一个最优解, 费用为$840(文件 toraEx5.2-2.txt). 下表描述

了这个最优解.

星期	磨锯片数 (当前日)			处理
	新锯片	通宵磨	2 天磨	
星期一	24(一)	10(二)+8(三)	6(四)	0
星期二	2(二)	6(三)	6(五)	0
星期三	0	14(四)	0	0
星期四	0	12(五)	8(日)	0
星期五	0	14(六)	0	4
星期六	0	14(日)	0	0
星期日	0	0	0	22

评注 表 5.13 中的模型仅适用于第一个星期的运行, 因为没考虑每星期每天轮流的特性, 即这星期的每一天可以作为下星期需求的起点. 处理这种情况的一种方法是假设第一个星期运行的每一天都使用新锯片. 这样, 对应于一星期的每一天, 可以形成一个正好由 7 个起点和 7 个终点组成的运输模型. 与表 5.13 相似, 只是在新模型去掉"新锯片"起点和"处理"终点. 还有, 只需要封住对角线上的元素 (单位费用 $= M$). 剩余的单元费用或者是 \$3.00 或者是 \$6.00. 例如对于 (星期六, 星期一) 的单元, 单位费用为 \$6.00; 对于单元 (星期六, 星期二)、(星期六, 星期三) 和 (星期六, 星期四) 和 (星期六, 星期五) 的单位费用为 \$3.00. 下表给出了费用为\$372 的解, 最优解总是要求只用 2 天的打磨服务. 本问题还存在其他一些最优解 (见文件 toraEx5.2-2a.txt).

第 i 周	第 $(i+1)$ 周							总计
	星期一	星期二	星期三	星期四	星期五	星期六	星期日	
星期一				6			18	24
星期二					8		4	12
星期三	12					2		14
星期四	8	12						20
星期五	4		14					18
星期六				14				14
星期日					10	12		22
总计	24	12	14	20	18	14	22	

习题 5.2A①

1. 在例 5.2-1 中, 假设每单位的储存费用是依据周期而变化的, 周期 1, 2, 3 的单位储存费用分别为 40, 30, 70 美分. 延期和生产费用按照该例中给出的保持不变. 确定最优解, 并解释其结果.

***2.** 在例 5.2-2 中, 假设锯片打磨服务提供了星期一和星期二 (第一天和第二天) 磨一个锯片\$1 的 3 天打磨服务. 重新建立模型, 并解释其最优解.

① 本习题中, 可以利用 TORA 程序求最优解. 运输问题的 AMPL 和 Excel 规划求解程序将在 5.3.2 节结尾部分介绍.

3. 在例 5.2-2 中, 如果一个锯片当天打磨过后没有使用, 则每个锯片每天的储存费用为 50 美分. 重新建立模型, 并解释其最优解.

4. JoShop 公司希望分配 4 种不同类型的机器来完成 5 种不同的加工任务. 能够使用的 4 种机器的数量分别为 25, 30, 20, 30 台, 5 种不同类型的加工任务量分别为 20, 20, 30, 10, 25 件. 第 4 种类型的机器不能分配给第 4 种类型的任务. 表 5.14 提供了分配每一种机器完成一类加工任务的单位费用 (用美元表示). 本问题的目标是确定分配给每种类型加工任务的每种类型机器的最优台数. 求解本问题, 并解释最优解.

表 5.14　第 4 题的单位费用表

		任务类型				
		1	2	3	4	5
机器类型	1	10	2	3	15	9
	2	5	10	15	2	4
	3	15	5	14	7	15
	4	20	15	13	—	8

***5.** 未来 4 个月对某种易过期物品的需求量分别为 400, 300, 420, 380 吨, 相应这 4 个月的供应能力为 500, 600, 200, 300 吨. 每个月每吨的采购费用不同, 分别为$100、$140、$120 和$150. 因为这种物品的易过期性, 当月生产的物品必须在 3 个月内 (包括生产月) 消费完. 每吨物品每月的储存费用为$3. 这种物品不能延期交货, 运用运输模型求解本问题, 并确定未来 4 个月的最优生产安排.

6. 一种小型发动机未来 5 个季度的需求量分别为 200, 150, 300, 250, 480 台, 生产商对应这 5 个季度的生产能力估计为 180, 230, 430, 300, 300 台. 不允许延期交货, 但在必要情况下, 生产商可安排加班满足当前需求. 每个周期的加班生产能力是正常生产能力的一半. 5 个季度的单位生产费用分别为$100、$96、$116、$102 和$106. 加班生产的产品的生产费用比正常生产费用高 50%. 如果现周期生产的发动机在以后周期使用, 则每台发动机每周期的额外储存费用为$4. 将本问题构造成一个运输模型, 并确定每个周期正常生产和加班生产发动机的数量.

7. 对飞机发动机要进行定期维修检查, 检查过程中有一种重要的部件必须更换. 以后 6 个月计划维修的飞机数量分别为 200, 180, 300, 198, 230, 290 架. 所有的维修工作必须在每月的第一天做完, 其间一个用过的部件可以更换为一个新部件, 或者更换为一个已检修过的部件. 用过的部件可以在当地工厂检修, 这些部件将准备为下个月的月初使用; 或送到中心修理厂, 这些部件将会延迟 3 个月使用 (包括维修月). 当地工厂的修理费为每个部件$120. 在中心工厂的修理费为每个部件$35. 要在下个月使用的已修好的部件每月每台的储存费用为 $1.5. 新部件在第一个月的采购价格为$200, 以后每两个月价格增加 5%. 用运输模型描述本问题, 并求出满足该零部件未来 6 个月需求的最优维修计划.

8. 国家公园服务中心收到 4 份要在阿肯色州的 3 个松木林伐木的投标书. 3 个林区的面积分别为 10 000, 20 000, 30 000 英亩. 单个投标人最多能够申请总英亩数的 50%. 3 个林区每英亩的投标价如表 5.15 所示. 投标人 2 不愿意投标林区 1, 投标人 3 不愿意投标林区 2.

(a) 在这种情况下, 服务中心希望总的投标收益最大. 说明如何把此问题表示一个运输模型.

(b) 确定 4 个投标人每人应得到的英亩数.

表 5.15 第 8 题每英亩投标价

		林地		
		1	2	3
投标人	1	\$520	\$210	\$570
	2	—	\$510	\$495
	3	\$650	—	\$240
	4	\$180	\$430	\$710

5.3 运输算法

运输算法完全采用单纯形算法 (第 3 章) 的步骤, 但运输算法没有采用常规单纯形表, 而是根据运输模型的特殊结构来构造一个更方便的计算方式.

特殊的运输算法是人们最早建立的手工计算的算法, 它充分考虑到计算的便捷. 现在, 有了强大的计算机程序, 可以把任何规模的运输模型化成一个 LP[①] 来解决. 然而除了它的历史意义之外, 还由于运输算法帮助我们从原理上深刻理解原始–对偶关系 (在 4.2 节介绍), 改进了它手工计算, 得到了实用的最终结果.

我们用下面的例子来解释算法的细节.

例 5.3-1 (SunRay 运输问题)

SunRay 运输公司从 3 个仓库把粮食运输到 4 个加工厂. 供应量 (按车数)、需求量 (也按车数), 以及不同路线每辆运粮车的单位运输费用如表 5.16 中的运输模型所示. 单位运输费用为 c_{ij}, (表示在每个方格的右上方) 以 100 美元为单位. 本模型寻找从仓库 i 到加工厂 $j(i = 1, 2, 3; j = 1, 2, 3, 4)$ 之间的最小费用运输计划量 x_{ij}.

运输算法 运输算法的步骤同单纯形算法完全一致.

第 1 步 确定一个初始的基本可行解, 转到第 2 步.

第 2 步 利用单纯形算法的最优性条件, 在所有非基变量中确定进基变量. 如果最优条件满足, 停止. 否则, 转到第 3 步.

第 3 步 利用单纯形算法的可行性条件, 在所有现有的基变量中确定离基变量, 寻找新的基本解. 返回到第 2 步.

---

① 事实上, TORA 程序利用标准的单纯性法在后台处理所有必要的计算过程, 采用运输模型的格式只是作为一种屏幕“展现”.

表 5.16 SunRay 的运输模型

	加工厂 1	加工厂 2	加工厂 3	加工厂 4	供应量
仓库 1	10 x_{11}	2 x_{12}	20 x_{13}	11 x_{14}	**15**
仓库 2	12 x_{21}	7 x_{22}	9 x_{23}	20 x_{24}	**25**
仓库 3	4 x_{31}	14 x_{32}	16 x_{33}	18 x_{34}	**10**
需求量	**5**	**15**	**15**	**15**	

5.3.1 初始解的确定

一个具有 m 个起点和 n 个终点的一般运输模型包含 $(m+n)$ 个约束方程, 每一个起点和终点都对应一个约束方程. 然而, 因为运输模型总是平衡的 (总供应量 = 总需求量), 这些约束方程中有一个方程是冗余的. 因此, 运输模型有 $(m+n-1)$ 个独立的约束方程, 即初始基本解由 $(m+n-1)$ 个基变量组成. 因此在例 5.3-1 中, 起始解有 $3+4-1=6$ 个基变量.

由于运输问题的特殊结构, 运用下面 3 种方法可以保证找到一个非人工的初始基本解[①]:

(1) 西北角法 (左上角法); (2) 最小费用法; (3) Vogel 近似法.

在初始解越好, 目标值越小的意义下, 这 3 种方法在产生的初始基本解的"质量"上有所差别. 一般来说, Vogel 近似法能够产生最好的初始基本解, 西北角法产生的初始解最差, 但同时, 西北角法所需要的计算量也最少.

西北角法 这个方法从表中西北角 (左上角) 的元素 (变量 x_{11}) 开始.

第 1 步 在表中尽量选择一个最大的元素, 使其相应的行和列都减去相应的选定元素的值.

第 2 步 删去零供应的行或零需求的列, 即表示不可能分配给那一行或那一列. 如果行和列都为零, 那么仅删去一个, 在没删去的行 (列) 中留下零供应 (需求).

第 3 步 如果恰好剩下一行或一列没有被删除, 停止. 否则, 如果列被删除则转移到右边, 如果行被删除则转移到下面. 转到第 1 步.

例 5.3-2

用上述算法求解例 5.3-1 的模型, 得到初始基本解如表 5.17 所示. 箭头表示分

① 这 3 种方法都有 TORA 教学程序模块, 见 5.3.3 节的结尾部分.

配运输量的生成顺序.

表 5.17 西北角法的初始解

	1	2	3	4	供应量
1	10 **5**	2 **10**	20	11	**15**
2	12	7 **5**	9 **15**	20 **5**	**25**
3	4	14	16	18 **10**	**10**
需求量	5	**15**	**15**	**15**	

初始基本解为

$$x_{11} = 5, \quad x_{12} = 10$$

$$x_{22} = 5, \quad x_{23} = 15, \quad x_{24} = 5$$

$$x_{34} = 10$$

这个计划相应的费用为

$$z = 5 \times 10 + 10 \times 2 + 5 \times 7 + 15 \times 9 + 5 \times 20 + 10 \times 18 = \$520$$

最小费用法 最小费用法是通过寻找最小费用的路径来找到一个更好的初始解. 这种方法尽量找到一个最小单位费用的单元. 然后, 删除满足条件的行或列, 调整相应的供应和需求量.

如果行和列都同时满足, 同西北角法一样, 仅需要删除一个. 重复检查没有被删去的行或列, 直到只剩下一行或一列为止.

例 5.3-3

用最小费用法, 按照下列步骤求解例 5.3-1:

(1) 在运输表中, 单元 (1,2) 有最小单位费用 (=\$2). 通过 (1,2) 最多能运输 $x_{12} = 15$ 辆运粮车, 这样刚好同时满足第 1 行和第 2 列, 删除第 2 列, 并把第 1 行的供应调整为 0.

(2) 单元 (3,1) 有最小未删除的单位费用 (=\$4). 选择 x_{31}=5, 因为第 1 列满足, 所以删除第 1 列, 调整第 3 行的需求为 10−5=5 辆运粮车.

(3) 继续采用同样的方法, 分配 15 辆运粮车到单元 (2,3), 0 辆运粮车到单元 (1, 4), 5 辆运粮车到单元 (3,4), 以及 10 辆运粮车到单元 (2,4)(请验证!)

得到的初始解如表 5.18 所示, 箭头表明了选定单元的顺序. 初始解 (由 6 个基变量组成) 是 $x_{12} = 15, x_{14} = 0, x_{23} = 15, x_{24} = 10$, $x_{31} = 5, x_{31} = 5$, $x_{34} = 5$, 相应的目标值为

$$z = 15 \times 2 + 0 \times 11 + 15 \times 9 + 10 \times 20 + 5 \times 4 + 5 \times 18 = \$475$$

表 5.18 最小费用法的初始解

	1	2	3	4	供应量
1	10	（开始） 2 **15**	20	11 **0**	**15**
2	12	7	9 **15**	（始束）20 **10**	**25**
3	4 **5**	14	16	18 **5**	**10**
需求量	**5**	**15**	**15**	**15**	

因为最小费用法产生了一个比西北角法 (例 5.3-2) 更小的 z 值 (\$475 比\$520), 所以最小费用法初始解的质量比西北角法好.

Vogel 近似法 (VAM) 一般来说 VAM 是最小费用法的改进版, 在大多数情况下, 产生的初始解比最小费用法更好.

第 1 步 对每一行 (列) 确定惩罚量: 在每一行 (列) 中找到一个最小的单位费用单元, 再在同一行 (列) 中找到一个次小的单位费用单元, 惩罚量即为次小的单位费用减去最小的单位费用.

第 2 步 找出惩罚量最大的行或列. 尽量分配给最小单位费用的单元最多的运粮车. 调整供应和需求, 删去已满足的行或列. 如果行和列同时满足, 删去二者之一, 分配给剩下的那个行 (列) 为零供应 (需求).

第 3 步 (a) 如果仅有一个未被删去的零供应的行或零需求的列, 则停止.

(b) 如果具有一个未被删去的大于零供应 (需求) 的行 (列), 那么采用最小费用方法确定行 (列) 的基变量. 停止.

(c) 如果所有的未被删去的行和列都有零供应和零需求, 那么采用最小费用方法确定零基变量. 停止.

(d) 否则, 转到第 1 步.

例 5.3-4

用 VAM 法求解例 5.3-1. 表 5.19 计算出了第一组惩罚量.

表 5.19 VAM 中的行惩罚和列惩罚

	1	2	3	4		行惩罚
1	10	2	20	11	**15**	10 − 2 = 8
2	12	7	9	20	**25**	9 − 7 = 2
3	4 **5**	14	16	18	**10**	14 − 4 = 10
	5	**15**	**15**	**15**		
列惩罚	10 − 4 = 6	7 − 2 = 5	16 − 9 = 7	18 − 11 = 7		

因为第 3 行有最大的惩罚 (=10), 这一行中, 单元 (3,1) 有最小的单位费用, 将 5 辆运粮车分配给 x_{31}. 第 1 列被满足, 必须删除掉. 接下来算出了新的惩罚量, 如表 5.20.

表 5.20 VAM 中的第一次赋值(x_{31}=5)

	1	2	3	4		行惩罚
1	**10**	2	20	11	**15**	9
2	**12**	7	9	20	**25**	2
3	**4** **5**	14	16	18	**10**	2
	5	**15**	**15**	**15**		
列惩罚	—	5	7	7		

表 5.20 表明第 1 行有最大的惩罚 (=9). 因此, 尽量分配最大的量给单元 (1,2), 使得 $x_{12}=15$, 第 1 行和第 2 列同时得到满足. 删去第 2 列, 调整第 1 行的供应量为零.

继续同样的办法, 第 2 行生成最大的惩罚 (=11), 分配 $x_{23}=15$, 删除第 3 列, 在第 2 行中留下 10 个单位. 仅留下第 4 列, 这列还有 15 个单位. 对该列采用最小费用法, 得到 $x_{14}=0, x_{34}=5$, $x_{24}=10$(请验证!). 这个解的相应目标值为

$$z=15\times 2+0\times 11+15\times 9+10\times 20+5\times 4+5\times 18=\$475$$

这个解恰好与最小费用法得到的目标值相同.

习题 5.3A

1. 比较下列每个模型中用西北角法、最小费用法和 Vogel 方法生成的初始解.

*(a)

0	2	1	**6**
2	1	5	**7**
2	4	3	**7**
5	**5**	**10**	

(b)

1	2	6	**7**
0	4	2	**12**
3	1	5	**11**
10	**10**	**10**	

(c)

5	1	8	**12**
2	4	0	**14**
3	6	7	**4**
9	**10**	**11**	

5.3.2 运输算法的迭代计算

(采用 5.3.1 节 3 种方法中的一种) 确定初始解后, 采用下面的算法来确定最优解:

第 1 步 采用单纯形法的*最优性条件*, 来确定能够改进解的作为当前非基变量的*进基变量*. 如果最优性条件满足, 停止. 否则, 转到第 2 步.

第 2 步 采用单纯形*可行性条件*确定*离基变量*. 改变基, 返回到第 1 步.

最优性条件和可行性条件不包括单纯形方法常用的行运算. 相反, 运输模型的特殊结构能够产生更简单的计算.

例 5.3-5

利用西北角法得到的解作为起点, 求解例 5.3-1 的运输模型.

表 5.21 给出西北角法产生的起始解, 求解方法如例 5.3-2 中的表 5.17.

从现有的非基变量 (不作为初始基本解的一部分) 中通过用**乘子法** (method of multiplier)(5.3.4 节介绍, 源于线性规划对偶理论) 计算 z 行的非基系数, 来求出进基变量.

在乘子法中, 用 u_i 和 v_j 来表示运输表中第 i 行和第 j 列的乘子. 对每一个现有的*基变量* x_{ij}, 5.3.4 节的这些乘子满足下面的方程:

$$u_i+v_j=c_{ij}, \text{对于每一个基变量} x_{ij}$$

表 5.21 开始迭代

	1	2	3	4	供应量
1	10 **5**	2 **10**	20	11	**15**
2	12	7 **5**	9 **15**	20 **5**	**25**
3	4	14	16	18 **10**	**10**
需求量	**5**	**15**	**15**	**15**	

如表 5.21 所示, 起始解有 6 个基变量, 得到 7 个未知数的 6 个方程. 为了解这些方程, 乘子法任意设置 $u_i = 0$, 然后求解其他变量, 如下表所示.

基变量	(u, v) 方程	解
x_{11}	$u_1 + v_1 = 10$	设置 $u_1 = 0 \to v_1 = 10$
x_{12}	$u_1 + v_2 = 2$	$u_1 = 0 \to v_2 = 2$
x_{22}	$u_2 + v_2 = 7$	$v_2 = 2 \to u_2 = 5$
x_{23}	$u_2 + v_3 = 9$	$u_2 = 5 \to v_3 = 4$
x_{24}	$u_2 + v_4 = 20$	$u_2 = 5 \to v_4 = 15$
x_{34}	$u_3 + v_4 = 18$	$v_4 = 15 \to u_3 = 3$

从上表可以得到

$$u_1 = 0, u_2 = 5, u_3 = 3$$
$$v_1 = 10, v_2 = 2, v_3 = 4, v_4 = 15$$

接下来, 使用 u_i 和 v_j, 通过对每个非基变量计算 $u_i + v_j - c_{ij}$ 的值, 对这些非基变量进行检查.

检查的结果如下表所示:

非基变量	$u_i + v_j - c_{ij}$
x_{13}	$u_1 + v_3 - c_{13} = 0 + 4 - 20 = -16$
x_{14}	$u_1 + v_4 - c_{14} = 0 + 15 - 11 = 4$
x_{21}	$u_2 + v_1 - c_{21} = 5 + 10 - 12 = 3$
x_{31}	$u_3 + v_1 - c_{31} = 3 + 10 - 4 = \mathbf{9}$
x_{32}	$u_3 + v_2 - c_{32} = 3 + 2 - 14 = -9$
x_{33}	$u_3 + v_3 - c_{33} = 3 + 4 - 16 = -9$

前面的信息加上对每个基变量 x_{ij} 都有 $u_i + v_j - c_{ij} = 0$ 的事实, 实际上相当于计算单纯形表的 z 行, 总的结果如下:

基	x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}	x_{31} ↓	x_{32}	x_{33}	x_{34}
z	0	0	−16	4	3	0	0	0	9	−9	−9	0

因为运输模型寻求费用*最小化*, 进基变量是具有 z 行中*最正*系数的变量. 因此, x_{31} 就是进基变量.

上述计算通常直接在运输表上进行, 如表 5.22 所示, 这表示其实并不需要直接写出 (u, v) 方程, 而是从设定 $u_1 = 0$ 开始①. 然后, 我们就可以计算在第 1 行中有基变量的所有列的 v 值, 即 v_1 和 v_2. 接下来, 根据基变量 x_{22} 的 (u, v) 方程计算 u_2, 有了 u_2, 我们计算 v_3 和 v_4. 最后, 用 x_{33} 的基本方程确定出 u_3. 当所有的 u 和 v 求出以后, 我们就能通过对每个非基变量 x_{ij} 计算 $u_i + v_j - c_{ij}$ 来检查这些非基变量, 检查过程如表 5.22 的左下角所示.

确定 x_{31} 为进基变量后, 我们还需要确定离基变量. 需要记住的是如果 x_{31} 进解成为基变量, 现有基变量中必须有一个离基成为 (零值) 非基变量.

选择 x_{31} 为进基变量意味着运输将采用这个路径, 因为它降低了总的运输费用. 通过这条新的路径我们最多能运输多少呢? 观察表 5.22 可知, 如果路径 (3,1) 运输 θ 单位 (即 $x_{31} = \theta$), 那么 θ 的最大值基于下面两个条件确定.

(1) 仍然满足供应上限和和需求约束.

(2) 通过所有路径的运输量仍然非负.

这两个条件决定了 θ 的最大值和离基变量. 首先, 建立一个起点和终点都在进基变量单元 (3, 1) 的*闭圈*(closed loop), 这个闭圈仅由*连接在一起*的水平线段和垂直线段组成 (不允许对角线)② , 除进基变量单元以外, 闭圈的每一个角都与一个基变量重合. 表 5.23 表示了 x_{31} 的闭圈. 一个闭圈仅对应于一个进基变量.

接下来, 分配 θ 给进基变量单元 (3, 1). 在供应和需求约束仍然满足的情况下, 如表 5.23 所示, 闭圈的每一个角在增加和减少 θ 值之间必须进行调整 (与闭圈的顺时针和逆时针无关). 对于 $\theta \geqslant 0$, 变量的新值仍然是非负的, 如果

$$x_{11} = 5 - \theta \geqslant 0$$

$$x_{22} = 5 - \theta \geqslant 0$$

$$x_{34} = 10 - \theta \geqslant 0$$

① TORA 指南显示, 对任何 u 或 v 指定一个零初始值并不影响最优解结果, 见 5.3.2 节末尾的 TORA 程序.

② TORA 指南模块可以让你交互式地确定闭圈单元, 并对你的选择是否正确会立即给出反馈信息, 见 5.3.2 节末尾的 TORA 程序.

表 5.22 第 1 步迭代的计算

	$v_1=10$	$v_2=2$	$v_3=4$	$v_4=15$	供应量
$u_1\equiv 0$	10 **5**	2 **10**	20 −16	11 4	**15**
$u_2=5$	12 3	7 **5**	9 **15**	20 **5**	**25**
$u_3=3$	4 9	14 −9	16 −9	18 **10**	**10**
需求量	**5**	**15**	**15**	**15**	

相应的 θ 的最大值是 5, 此时 x_{11} 和 x_{22} 达到了零点. 因为仅有一个现有基变量离开基本解, 我们可选择 x_{11} 或 x_{22} 作为离基变量. 随便选择 x_{11} 离开基本解.

表 5.23 x_{31} 闭圈的确定

	$v_1=10$	$v_2=2$	$v_3=4$	$v_4=15$	供应量
$u_1\equiv 0$	10 **5 − Θ** −	2 **10 + Θ** +	20 −16	11 4	**15**
$u_2=5$	12 3	7 **5 − Θ** −	9 **15**	20 **5 + Θ** +	**25**
$u_3=3$	4 Θ + 9	14 −9	16 −9	18 **10 − Θ** −	**10**
需求量	**5**	**15**	**15**	**15**	

选择 $x_{31}(=5)$ 进基和 x_{11} 离基, 这需要调整表 5.24 中所显示的闭圈的角上的基变量的值. 因为通过路径 (3, 1) 运输的每单位运输费用降低了 $\$9(=u_3+v_1-c_{31})$, 对应于新计划的总费用比以前计划的总费用降低了 $\$9\times 5=\$45$. 因此, 新的费用为 $\$520-\$45=\$475$.

已知新的基本解后, 我们重复乘子 u 和 v 的计算过程, 如表 5.24 所示. 进基变量是 x_{14}, 闭圈表明 $x_{14}=10$, 离基变量为 x_{24}.

如表 5.25 所示, 产生的新解比上一个解减少费用 $\$4\times 10=\$40$, 因此新的费用为 $\$475-\$40=\$435$, 新的 $u_i+v_j-c_{ij}$ 对所有的非基 x_{ij} 都是非负的了. 因此, 表 5.25 中的解是最优的.

下表给出了最优解.

表 5.24　第 2 步迭代计算

	$v_1=1$	$v_2=2$	$v_3=4$	$v_4=15$	供应量
$u_1\equiv 0$	10 −9	2 **15 − Θ** −	20 −16	11 Θ +　4	**15**
$u_2=5$	12 −6	7 **0 + Θ** +	9 **15**	20 **10 − Θ** −	**25**
$u_3=3$	4 **5**	14 −9	16 −9	18 **5**	**10**
需求量	**5**	**15**	**15**	**15**	

表 5.25　第 3 步迭代计算(最优)

	$v_1=-3$	$v_2=2$	$v_3=4$	$v_4=11$	供应量
$u_1\equiv 0$	10 −13	2 **5**	20 −16	11 **10**	**15**
$u_2=5$	12 −10	7 **10**	9 **15**	20 −4	**25**
$u_3=7$	4 **5**	14 −5	16 −5	18 **5**	**10**
需求量	**5**	**15**	**15**	**15**	

从仓库	到工厂	运输车数
1	2	5
1	4	10
2	2	10
2	3	15
3	1	5
3	4	5
	最优费用 =$435	

TORA 程序

从 Solve/Modify Menu 中, 选择 Solve ⇒ Iterations, 并选择 3 种方法 (西北角、最小费用和 Vogel) 之一开始运输模型的迭代. 迭代模块提供了两种有用的交互特征:

(1) 在生成迭代 2 之前 (默认情况是 $u_1=0$), 可以设置任意 u 或 v 为零. 观察

可知, 虽然 u_i 和 v_j 的值发生了改变, 非基单元的检查 $(= u_i + v_j - c_{ij})$ 仍然不变. 这就表明开始时任何 u 或 v 可以设置为零 (实际上, 任何值), 而不会影响最优值的计算.

(2) 通过 (以任何顺序) 单击组成路径的角单元, 你可以测试闭圈选择的正确性. 如果你的选择是正确的, 单元的颜色将会改变 (进基变量为绿色, 离基变量为红色, 其他颜色为灰色)

Excel 规划求解程序

把运输模型直接输入到 Excel 电子表格. 图 5.4 提供了例 5.3-1 的 Excel 规划求解程序模板 (文件 solverEx5.3-1.xls), 以及所有涉及的公式和定义.

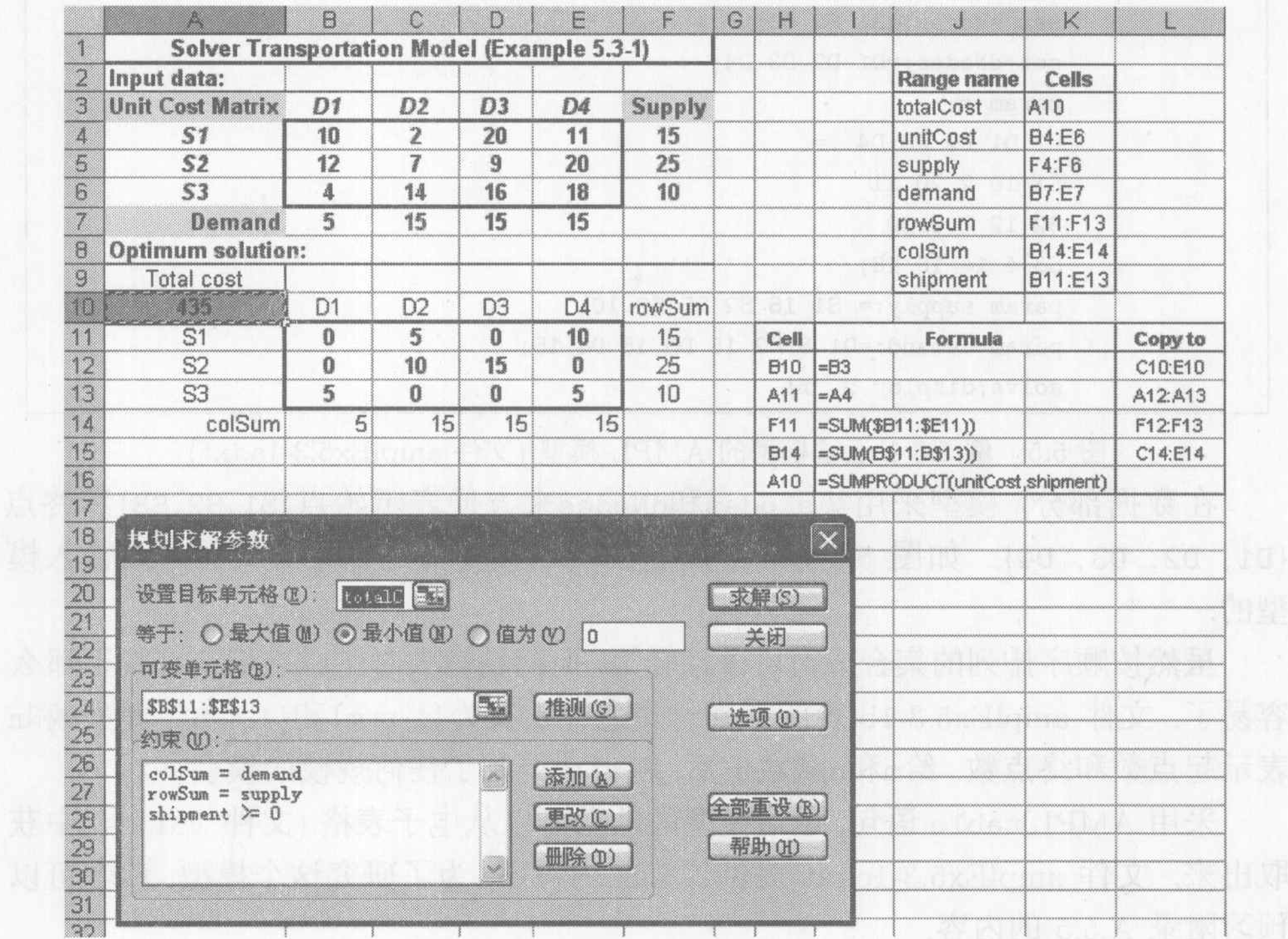


图 5.4　例 5.3-1 运输模型的 Excel 规划求解 (文件 solver Ex5.3-1.xls)

在输入部分, 数据包括单位费用矩阵 (单元 B4:E6)、起点名称 (单元 A4:A6)、终点名称 (单元 B3:E3)、供应量 (单元 F4:F6) 和需求量 (单元 B7:E7). 在输出部分, 单元 B11:E13 提供了矩阵格式的最优解. 总费用公式在目标单元 A10 中.

AMPL 程序

图 5.5 提供了例 5.3-1 运输模型的 AMPL 模型 (文件 amplEx5.3-1.txt). 模型中用到的名字都是自我解释型的. 约束和目标函数都采用例 5.1-1 中线性规划模型的格式.

```
#----- Transporation model (Example 5.3-1)-----
set sNodes;
set dNodes;
param c{sNodes,dNodes};
param supply{sNodes};
param demand{dNodes};
var x{sNodes,dNodes}>=0;
minimize z:sum {i in sNodes,j in dNodes}c[i,j]*x[i,j];
subject to
source {i in sNodes}:sum{j in dNodes}x[i,j]=supply[i];
dest {j in dNodes}:sum{i in sNodes}x[i,j]=demand[j];
data;
set sNodes:=S1 S2 S3;
set dNodes:=D1 D2 D3 D4;
param c:
   D1 D2 D3 D4 :=
S1 10 2 20 11
S2 12 7 9 20
S3 4 14 16 18;
param supply:= S1 15 S2 25 S3 10;
param demand:=D1 5 D2 15 D3 15 D4 15;
solve;display z,x;
```

图 5.5　例 5.3-1 运输模型的 AMPL 模型 (文件 amplEx5.3-1a.txt)

在数据部分, 模型采用集`sNodes`和`dNodes`来方便表示始点`{S1,S2,S3}`和终点`{D1, D2, D3, D4}`. 如图 5.5 所示, 所有的输入数据都是以集成员的方式输入模型的.

虽然按顺序排列的集合成员可读性较好, 但问题较大时生成它们或许就不那么容易了. 文件 amplEx5.3-1b 表明了同一集怎样定义为`{1···m}`和`{1···n}`, 此处`m`和`n`表示起点数和终点数. 给`m`和`n`赋数值后, 就自动得到了任何规模的模型.

采用 AMPL `table` 语句, 运输模型的数据可以从电子表格 (文件 TM.xls) 中获取出来. 文件 amplEx5.3-1c.txt 提供了相应的程序. 为了研究这个模型, 读者可以预习附录 A.5.5 的内容.

习题 5.3B

1. 考虑表 5.26 中的运输模型.

表 5.26　第 1 题的运输模型

(i)				(ii)				(iii)			
\$0	\$2	\$1	6	\$10	\$4	\$2	8	–	\$3	\$5	4
\$2	\$1	\$5	9	\$2	\$3	\$4	5	\$7	\$4	\$9	7
\$2	\$4	\$3	5	\$1	\$2	\$0	6	\$1	\$8	\$6	19
5	5	10		7	6	6		5	6	19	

(a) 用西北角法寻找初始解.

(b) 设计求最优解的迭代步骤.

(c) TORA实验: 用 TORA 的迭代模块, 比较西北角法、最小费用法和 Vogel 法求最优解的迭代效果.

(d) Excel 规划求解实验: 修改文件 solverEx5.3-1.xls 解决此问题.

(e) AMPL实验: 修改文件 amplEx5.3-1b.xls 解决此问题.

2. 在表 5.27 的运输问题中, 总需求超过总供应. 假设对终点 1, 2, 3 的未满足需求部分的每单位惩罚费用分别为\$5、\$3 和\$2. 用最小费用法的初始解, 并迭代计算出最优解.

表 5.27 第 2 题的数据

\$5	\$1	\$7	**10**
\$6	\$4	\$6	**80**
\$3	\$2	\$5	**15**
75	**20**	**50**	

3. 在第 2 题中, 假定没有惩罚费用, 但终点 3 的需求必须全部满足.

(a) 寻找最优解.

(b) Excel 规划求解实验: 修改文件 solverEx5.3-1.xls 求解此问题.

(c) AMPL实验: 修改文件 amplEx5.3b-1.xls 求解此问题.

4. 表 5.28 是一个不平衡的运输问题, 假如起点有一个单位没有运出 (到任何终点), 起点 1, 2, 3 的储存费用分别为\$5、\$4 和\$3. 另外为了给新的产品腾出空间, 起点 2 的供应量必须全部运出. 运用 Vogel 法计算初始解, 并给出寻找最优解的迭代.

表 5.28 第 4 题的数据

\$1	\$2	\$1	**20**
\$3	\$4	\$5	**40**
\$2	\$3	\$3	**30**
30	**20**	**20**	

5. 在一个 3×3 的运输问题中, 令 x_{ij} 表示从起点 i 到终点 j 的运输量, c_{ij} 表示相应的单位运输费用. 起点 1, 2, 3 的供应量分别为 15, 30, 85 单位, 终点 1, 2, 3 的需求量分别为 20, 30, 80 单位. 假设用西北角法求出的初始解是最优解, 相应的乘子为 $u_1 = -2, u_2 = 3, u_3 = 5, v_1 = 2, v_2 = 5$ 和 $v_3 = 10$.

(a) 寻找相应的最优费用.

(b) 保持西北角解的最优性, 对每一个非基变量求最小的 c_{ij}.

6. 表 5.29 中的运输问题给出了退化的基本解 (即至少有一个基变量为零). 假设对应解的乘子是 $u_1 = 1, u_2 = -1, v_1 = 2, v_2 = 2, v_3 = 5$, 所有 (基本和非基本) 零 x_{ij} 变量的单位费用为

$$c_{ij} = i + j\theta, \quad -\infty < \theta < \infty$$

(a) 若给定的解是最优的, 确定目标函数的最优值.

(b) 求能够保证给定解的最优性的 θ 值. (提示: 找到零基变量的位置)

表 5.29 第 6 题的数据

10			10
	20	20	40
10	20	20	

7. 考虑问题:

$$\min \quad z=\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} x_{ij} \geqslant a_i, \quad i=1,2,\ldots,m$$

$$\sum_{i=1}^{m} x_{ij} \geqslant b_j, \quad j=1,2,\ldots,n$$

$$x_{ij} \geqslant 0, \quad 对于所有的 i,j$$

似乎可以认为, 若 $\sum a_i \geqslant \sum b_j (\sum a_i \leqslant \sum b_j)$, 最优解将会要求第一组 (第二组) 不等式用等式约束代替. 但表 5.30 的反例表明, 这个假设是不正确的.

表 5.30 第 7 题的数据

\$1	\$1	\$2	5
\$6	\$5	\$1	6
2	7	1	

证明用这种方法可得到解 $x_{11}=2, x_{12}=3, x_{22}=4, x_{23}=2$, $z=\$27$, 这比可行解 $x_{11}=2, x_{12}=7, x_{23}=6$, $z=\$15$ 还要差.

5.3.3 乘子法的单纯形方法解释

乘子法与单纯形方法之间的关系可用原始–对偶关系 (见 4.2 节) 来解释. 根据表示运输模型的线性规划的特殊结构 (见例 5.1-1 的解释), 相应的对偶问题可以写为

$$\max \ z=\sum_{i=1}^{m} a_i u_i+\sum_{j=1}^{n} b_j v_j$$

$$\text{s.t.} \quad u_i+v_j \leqslant c_{ij}, 对所有的\ i\ 和\ j$$

$$u_i\ 和\ v_j\ 无符号限制$$

其中

$$a_i = 起点\ i\ 的供应量$$

$$b_j = 终点\ j\ 的需求量$$

$$c_{ij} = \text{从起点 } i \text{ 到终点 } j \text{ 的单位运输费用}$$
$$u_i = \text{关于起点 } i \text{ 约束条件的对偶变量}$$
$$v_j = \text{关于终点 } j \text{ 约束条件的对偶变量}$$

根据 4.2.4 节的公式 2, 变量 x_{ij} 的目标函数系数 (简约费用) 等于相应对偶约束的左端项与右端项之差, 即 $u_i + v_j - c_{ij}$. 但是我们知道, 对每个基变量, 这个量一定等于零, 这样就得到了下面的结果:

$$u_i + v_j = c_{ij}, \quad \text{对每一个基变量} x_{ij}.$$

共有 $(m+n-1)$ 个这样的方程, 这些方程的解 (在随便设定一个值 $u_1 = 0$ 后) 可以得到乘子 u_i 和 v_j. 计算得到这些乘子以后, 我们就能从所有非基变量中选择一个具有最大正数的 $u_i + v_j - c_{ij}$ 作为进基变量.

对一个对偶变量分配一个任意值 (如 $u_1 = 0$) 看起来可能会与在 4.2.3 节采用方法 2 计算对偶变量的办法不一样. 实际上, 对于一个给定的基本解, 对偶值一定是唯一的 (反之亦然). 关于这一点见习题 5.3A 的第 2 题.

习题 5.3C

1. 写出例 5.3-5(表 5.21) 中运输问题的线性规划的对偶问题. 利用表 5.25 中的最优对偶值计算相应的最优对偶目标值, 并证明它等于该例子给出的最优费用.
2. 在运输模型中, 其中一个对偶变量赋予了一个任意值. 这意味着对于同一个基本解, 相应对偶变量的值是不唯一的. 这个结果似乎与线性规划理论相矛盾, 其中对偶值等于基变量的目标系数向量与相应基的逆矩阵的乘积 (见方法 2, 4.2.3 节). 证明对于运输模型, 虽然基矩阵的逆是唯一的, 但基变量目标系数向量并不一定唯一. 特别地, 证明对所有 i 和 j, 如果 c_{ij} 变成 $c_{ij} + k$, 这里 k 是一个常数, 那么 x_{ij} 的最优值将保持不变. 因此, 给一个对偶变量赋任意值就等价于所有的 c_{ij} 加上一个常数 k.

5.4 指 派 模 型

"派合适的人做合适的事", 这是对指派模型最贴切的描述. 指派模型可以用对不同技术水平的工人分派工作来解释. 不同的工人完成不同的工作的费用是不一样的, 目标是确定指派工人到工作岗位的最小指派费用.

表 5.31 表示了 n 个工人和 n 件工作的指派模型.

元素 c_{ij} 表示指派工人 i 做工作 $j(i, j = 1, 2, \cdots, n)$ 的费用. 不失一般性, 假设工人的人数等于工作的个数, 因为我们总可以添加一个虚设的工作, 或虚设的工人来使得这一假设得到满足.

指派模型实际上是运输模型的一种特殊情况, 工人可表示为起点, 工作可表示为终点. 在每一起点 (终点) 的供应 (需求) 量正好等于 1. 工人 i 到工作 j 的“运输”费用是 c_{ij}. 实际上, 指派模型可以直接转化为常规运输模型来解决. 然而, 供应量和需求量都等于 1 这一事实可以让我们利用**匈牙利算法** (Hungarian method)[①] 来方便地求出指派模型的解. 虽然新的解法与运输模型无关, 但这个算法如同运输模型一样, 也是基于单纯形方法.

表 5.31　指派模型

		工作				
		1	2	…	n	
工人	1	c_{11}	c_{12}	…	c_{1n}	1
	2	c_{21}	c_{22}	…	c_{2n}	1
	⋮	⋮	⋮		⋮	⋮
	n	c_{n1}	c_{n2}	…	c_{nn}	1
		1	1	…	1	

5.4.1　匈牙利算法

我们先用两个例子来说明这一新算法的机制. 5.4.2 节给出对该方法的一种基于单纯形法的解释.

例 5.4-1

Joe Klyne 的 3 个孩子 John、Karen 和 Terri 想挣点钱, 用作学校组织去当地动物园时的零花钱. Klyne 先生为孩子们选择了 3 样杂活: 锄草、粉刷车库门和洗车. 为了避免孩子们之间的争抢, Klyne 先生要求孩子们 (秘密地) 提交投标方案, 说明对每样杂活他们认为该付多少钱合适. 我们可以理解为, 这 3 个孩子都会服从爸爸的决定, 由爸爸分配哪个孩子干哪件家务. 表 5.32 表示了收到的投标方案. 基于这些信息, Klyne 先生该怎样分配这些活呢?

表 5.32　Klyne 的指派问题

	除草	粉刷	洗车
John	\$15	\$10	\$9
Karen	\$9	\$15	\$10
Terri	\$10	\$12	\$8

可以利用匈牙利法来求解这个指派问题.

第 1 步　对于初始的费用矩阵, 找出每行最小值的元素, 本行所有元素减去这个最小值.

① 之所以归为运输模型, 是因为匈牙利法最初的设计源于手工计算的方便, 放在本节介绍完全是基于历史的原因. 现在这种计算运用计算机程序设计的 LP 来解决问题更为有效.

第 2 步 对于第 1 步得到的矩阵, 找出每列最小值的元素, 本列所有元素减去这个最小值.

第 3 步 利用第 2 步得到的矩阵, 由其值为零的元素确定最优解.

设 p_i 和 q_j 分别表示第 1 步和第 2 步中确定的 i 行和 j 列相应的最小费用. 表 5.33 表示了初始的费用矩阵和第 1 步计算得到的每行的最小值.

表 5.33 匈牙利法的第 1 步

	除草	粉刷	洗车	行最小值
John	15	10	9	$p_1 = 9$
Karen	9	15	10	$p_2 = 9$
Terri	10	12	8	$p_3 = 8$

接下来, 表 5.34 表示了每一行的每一元素减去行最小值后得到的简化矩阵.

表 5.34 匈牙利法的第 2 步

	除草	粉刷	洗车
John	6	1	0
Karen	0	6	1
Terri	2	4	0
列最小值	$q_1 = 0$	$q_2 = 1$	$q_3 = 0$

第 2 步产生了每列的最小值, 如表 5.34 所示. 每列每一元素分别减去列最小值, 得到简化矩阵如表 5.35 所示. 带有下划线的值为零的单元给出了最优解. 即 John 得到了粉刷车库门的工作, Karen 得到了锄草的工作, Terri 去洗车. Klyne 先生支出的总费用是 9+10+8=\$27. 这笔钱又总是等于 $(p_1+p_2+p_3)+(q_1+q_2+q_3) = (9+9+8)+(0+1+0) = \$27$. (这个结果的证明在下节给出.)

表 5.35 匈牙利法的第 3 步

	除草	粉刷	洗车
John	6	$\underline{0}$	0
Karen	$\underline{0}$	5	1
Terri	2	3	$\underline{0}$

在上例中, 因为最后得到的矩阵中值为零的单元恰好能够产生一个可行的指派方案 (即每一个孩子分到了一份具体的工作), 匈牙利法给出的步骤运行正常. 但在一些情况下, 由第 1 步和第 2 步产生的零值单元不一定能直接得到一个可行解, 还需要更多的步骤得到最优 (可行) 指派. 下面的例子表示了这种情况.

例 5.4-2

假设把例 5.4-1 的问题扩展到有 4 个孩子和 4 项杂活, 表 5.36 表示了问题的

费用项.

表 5.36　指派问题

		杂活 1	2	3	4
孩子	1	\$1	\$4	\$6	\$3
	2	\$9	\$7	\$10	\$9
	3	\$4	\$5	\$11	\$7
	4	\$8	\$7	\$8	\$5

用第 1 步和第 2 步处理表 5.36(用 $p_1 = 1, p_2 = 7, p_3 = 4, p_4 = 5, q_1 = 0, q_2 = 0, q_3 = 3, q_4 = 0$), 得到如表 5.37 所示的简化矩阵.(请验证!)

表 5.37　简约指派矩阵

		杂活 1	2	3	4
孩子	1	0	3	2	2
	2	2	0	0	2
	3	0	1	4	3
	4	3	2	0	0

零单元的位置并不能把这些单项工作分派给所有的孩子. 例如, 如果把杂活 1 分派给孩子 1, 那么, 第 1 列将会消失, 孩子 3 在剩下的 3 列中找不到值为零的单元. 这个障碍可以通过在例 5.4-1 所述方法中添加下面的一步来克服.

第 2a 步　假如通过第 1 步和第 2 步不能保证得到可行的指派 (在所有的零单元内),

(i) 在最后的简化矩阵中画出最少的垂直和水平线, 覆盖所有零单元.

(ii) 在没覆盖的单元中, 选择值最小的单元, 每一个没覆盖的单元都减去这个最小值. 然后, 在两条线的交叉点处都加上这个最小值.

(iii) 在得到的值为零的单元中如果没有找到可行的指派, 重复第 2a 步. 否则, 转到第 3 步, 决定最优指派.

把第 2a 步应用到最后的矩阵, 得到表 5.38 中的阴影单元. 没有阴影覆盖的单元的最小值 (斜体显示) 等于 1. 这个单元的值加到粗体交叉单元中, 从而得到表 5.39 的矩阵.

最优解 (由加了下划线的零表示) 表明：孩子 1 干杂活 1, 孩子 2 干杂活 3, 孩子 3 干杂活 2, 孩子 4 干杂活 4. 相应的最优费用是 $1 + 10 + 5 + 5 = \$21$. 累加 p_i 和 q_j 可以得到相同的费用, 在阴影单元确定以后, 单元被减去的值为 $(1 + 7 + 4 + 5) + (0 + 0 + 3 + 0) + (1) = \$21$.

表 5.38 应用第 2a 步

		杂活 1	2	3	4
孩子	1	0	3	2	2
	2	2	0	0	2
	3	0	1	4	3
	4	3	2	0	0

表 5.39 最优指派

		杂活 1	2	3	4
孩子	1	**0**	2	1	1
	2	3	0	**0**	2
	3	0	**0**	3	2
	4	4	2	0	**0**

AMPL 程序

文件 amplEx5.4-2.txt 对指派模型提供了 AMPL 模型, 它和运输模型非常相似.

习题 5.4A

1. 解表 5.40 中的指派模型.
   (a) 用匈牙利法求解.
   (b) TORA实验: 把问题表示成线性规划, 用 TORA 求解.
   (c) TORA实验: 把问题表示成运输问题, 用 TORA 求解.
   (d) Excel 规划求解实验: 修改 Excel 文件 solverEx5.3-1.xls 求解此问题.
   (e) AMPL实验: 修改 amplEx5.3-1 解决此问题.

表 5.40 第 1 题的数据

(i)

\$3	\$8	\$2	\$10	\$3
\$8	\$7	\$2	\$9	\$7
\$6	\$4	\$2	\$7	\$5
\$8	\$4	\$2	\$3	\$5
\$9	\$10	\$6	\$9	\$10

(ii)

\$3	\$9	\$2	\$3	\$7
\$6	\$1	\$5	\$6	\$6
\$9	\$4	\$7	\$10	\$3
\$2	\$5	\$4	\$2	\$1
\$9	\$6	\$2	\$4	\$5

2. JoShop 公司需要把 4 项工作派给 4 个工人. 完成一项工作的费用是工人技能的函数. 表 5.41 是这些指派的费用. 工人 1 不能做工作 3, 工人 3 不能做工作 4. 运用匈牙利法求出最优指派方案.

表 5.41　第 2 题的数据

		工作			
		1	2	3	4
工人	1	\$50	\$50	—	\$20
	2	\$70	\$40	\$20	\$30
	3	\$90	\$30	\$50	—
	4	\$70	\$20	\$60	\$70

3. 在第 2 题的 JoShop 模型中. 假设有另外一个工人 (第 5 个) 能完成所有这 4 项工作, 完成每项工作相应费用为\$60、\$45、\$30 和\$80. 把原 4 个工人中的任何一个换成这个新工人划算吗?

4. 在第 2 题的模型中, 假设 JoShop 有了第 5 项工作, 目前 4 个工人完成这项工作的费用分别为\$20、\$10、\$20 和\$80. 这项新工作比 JoShop 的原有的 4 项工作都优先吗?

***5.** 一位经理需要在达拉斯总部和亚特兰大分部之间做 4 次往返旅行, 如表 5.42 所示.

表 5.42　第 5 题的数据

从达拉斯出发日期	返回达拉斯的日期
6 月 3 日, 星期一	6 月 7 日, 星期五
6 月 10 日, 星期一	6 月 12 日, 星期三
6 月 17 日, 星期一	6 月 21 日, 星期五
6 月 25 日, 星期二	6 月 28 日, 星期五

从达拉斯出发的往返旅行机票的价格为\$400. 如果一张机票的到达和出发日期跨越周末 (星期六和星期天) 的话, 可以打 25%的折扣. 若在亚特兰大停留 21 天, 折扣率增加到 30%. 一张达拉斯与亚特兰大之间 (两个方向都可以) 的单程票价为\$250. 这位经理该怎样买票呢?

***6.** 图 5.6 是某个机械加工车间的布局图, 小方块 1, 2, 3, 4 表示现有的工作中心, 而小圆圈 a, b, c, d 表示将要新增加的 4 个工作中心 I, II, III, IV. 问题的目标是指派这些新中心到所建议的地点, 使得在现有的中心和所建议的中心之间材料传输的总距离最短. 表 5.43 表示新中心和现有中心之间的传输频率. 原材料传输设备沿着交叉在中心的矩形通道行走. 例如, 中心 1 和 b 点单程的行走距离 (以米为单位计) 是 $30 + 20 = 50$ m.

7. 在阿肯色大学的工业工程系, INEG 4904 是一门高级设计课程, 目的是让学生以设计小组的形式运用本科课程学到的知识和技能解决实际问题. 每个小组的成员选出一位项目经理, 提出适当的研究范围, 撰写和提交项目方案, 完成项目目标规定的任务, 并撰写和提交最终项目报告. 课程指导教师给出可能的工程项目, 对每一工程提供适当的信息表, 包括与主办单位的联系、工程描述, 以及完成工程所需要的技能. 课程要求每个设计小组提交一份报告,

说明如何选择小组成员和经理. 这份报告还要按照优先级对每个项目排定顺序, 包括说明小组的技能与项目目标的符合程度. 在某个指定的学期, 提出了以下工程项目: Boeing F-15, Boeing F-18, Boeing Simulation, Cargil, Cobb-Vantress, ConAgra, Cooper, DaySpring (规划), DaySpring (物资处理), J. B. Hunt, Raytheon, Tyson South, Tyson East, Wal-Mart 和 Yellow Transportation. Boeing 和 Raytheon 项目要求, 所有项目小组成员必须是美国公民. 在这个学期的 11 个设计小组中, 有 4 个小组没有符合这项要求.

表 5.43 第 6 题的数据

		新中心			
		Ⅰ	Ⅱ	Ⅲ	Ⅳ
	1	10	2	4	3
现有	2	7	1	9	5
中心	3	0	8	6	2
	4	11	4	0	7

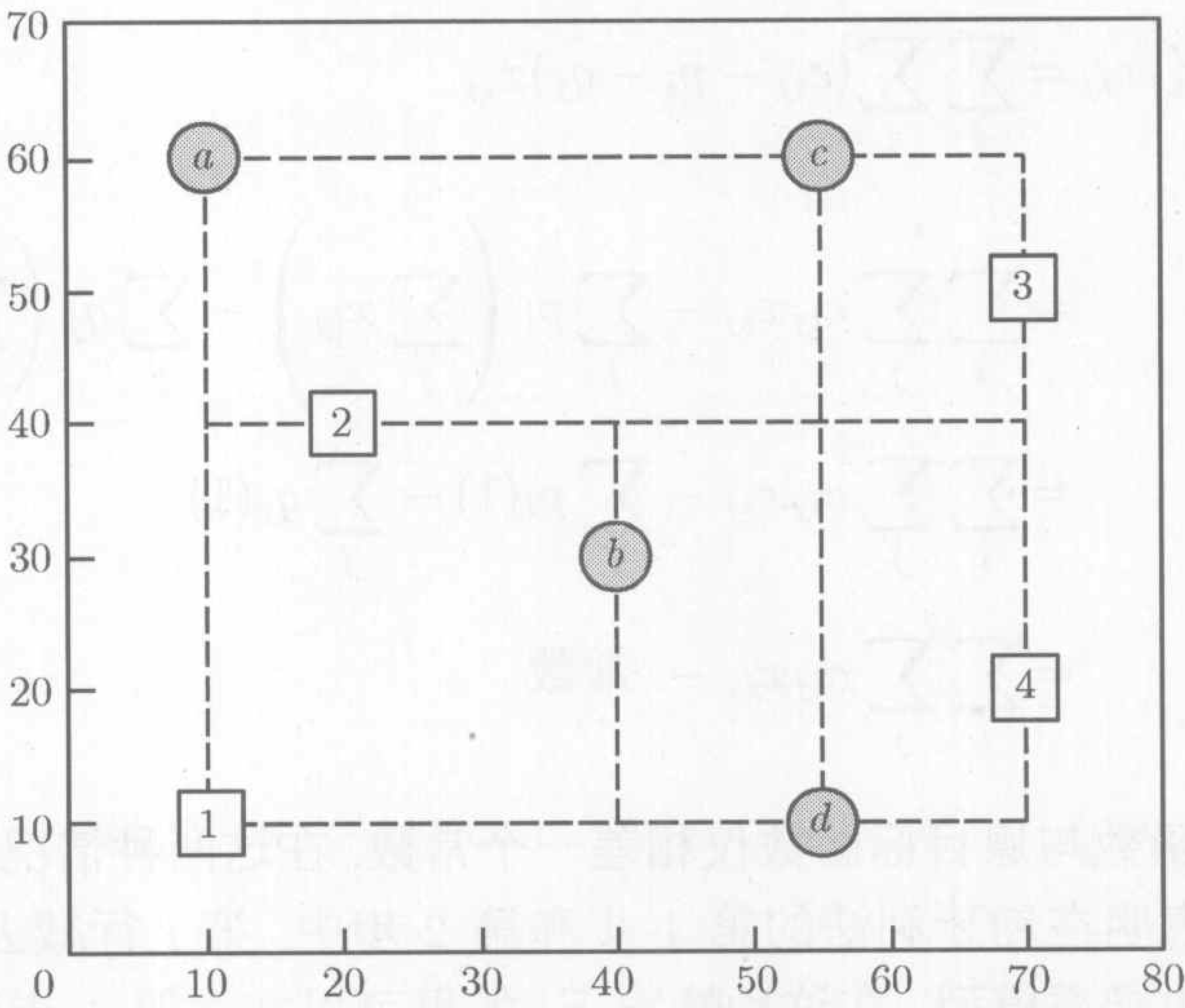


图 5.6 习题 5.4A 第 6 题的机器车间的布局图

设计一套方案, 把这些工程项目指派到各个小组, 并说明你的决策理由.

5.4.2 匈牙利算法的单纯形解释

指派 n 个工人完成 n 项工作的指派模型可以按下列方式表述为线性规划问题: 令 c_{ij} 表示指派工人 i 完成工作 j 的费用, 并定义

$$x_{ij} = \begin{cases} 1, & \text{若指派工人 } i \text{ 完成工作 } j \\ 0, & \text{若没有指派工人 } i \text{ 完成工作 } j \end{cases}$$

则给出单纯形模型如下:

$$\min\ z=\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}x_{ij}$$

$$\text{s.t.}\ \sum_{j=1}^{n}x_{ij}=1,\quad i=1,2,\cdots,n$$

$$\sum_{i=1}^{n}x_{ij}=1,\quad j=1,2,\cdots,n$$

$$x_{ij}=0\ 或\ 1$$

当费用矩阵 (c_{ij}) 任一行或列增加或减少一个常数时, 上述线性规划的最优解不变. 为了证明这一点, 令 p_i 和 q_j 为从第 i 行和第 j 列减去的常数, 费用元素 c_{ij} 变为

$$c'_{ij}=c_{ij}-p_i-q_j$$

那么

$$\begin{aligned}\sum_i\sum_j c'_{ij}x_{ij}&=\sum_i\sum_j(c_{ij}-p_i-q_j)x_{ij}\\&=\sum_i\sum_j c_{ij}x_{ij}-\sum_i p_i\left(\sum_j x_{ij}\right)-\sum_j q_j\left(\sum_i x_{ij}\right)\\&=\sum_i\sum_j c_{ij}x_{ij}-\sum_i p_i(1)-\sum_j q_j(1)\\&=\sum_i\sum_j c_{ij}x_{ij}-\ 常数\end{aligned}$$

因为新目标函数与原目标函数仅相差一个常数, 在这两种情况下, 最优解的 x_{ij} 值是一样的. 这表明在匈牙利法的第 1 步和第 2 步中, 第 i 行减去 p_i 和第 j 列减去 q_j, 产生等价的指派模型. 在这种情况下, 如果可以在由第 1 步和第 2 步产生的费用矩阵中的 n 个值为零的单元中发现一个可行解, 那么, 它一定是一个最优解, 因为在修正矩阵中费用不能小于零.

如果生成的 n 个值为零的单元不能得到一个可行解 (如例 5.4-2 所示), 则应用第 2a 步 (处理零单元的覆盖). 这个方法的正确性由线性规划单纯形法保证, 可以由对偶理论 (第 4 章) 和互补松弛定理 (第 13 章, 见本书下册) 来解释. 因为那些章节已经有一些叙述, 这里不再给出详细的证明.

$(p_1+p_2+\cdots+p_n)+(q_1+q_2+\cdots+q_n)$ 给出最优目标值的理由是, 它代表了指派模型的对偶目标函数. 这个结果可通过与在 5.3.3 节给出的运输模型的对偶目标函数的比较中看出. [见 Bazaraa 等人 (1990, pp. 499−508) 的内容.]

5.5 转 运 模 型

转运模型指货物在达到最后的终点之前通过中间或过渡节点来运输，认为这样比直接把货物运到最后终点或许会节省费用. 这个概念比通常的运输模型的概念更概括，而通常的运输模型是直接把货物从起点运送到终点.

本节说明如何使用**缓冲点**(buffer) 思想把转运模型变换成常规的运输模型.

例 5.5-1

两个汽车制造厂 P1 和 P2，通过两个转运中心 T1 和 T2，供应 3 个经销商 D1，D2，D3，如图 5.7 所示. 制造厂 P1 和 P2 的供应量分别为 1 000 和 1 200 辆汽车，经销商 D1，D2，D3 的需求量分别为 800，900，500 辆汽车. 网络的连接线 (弧) 上显示了每两个节点之间的单位运输费用 (单位：100 美元).

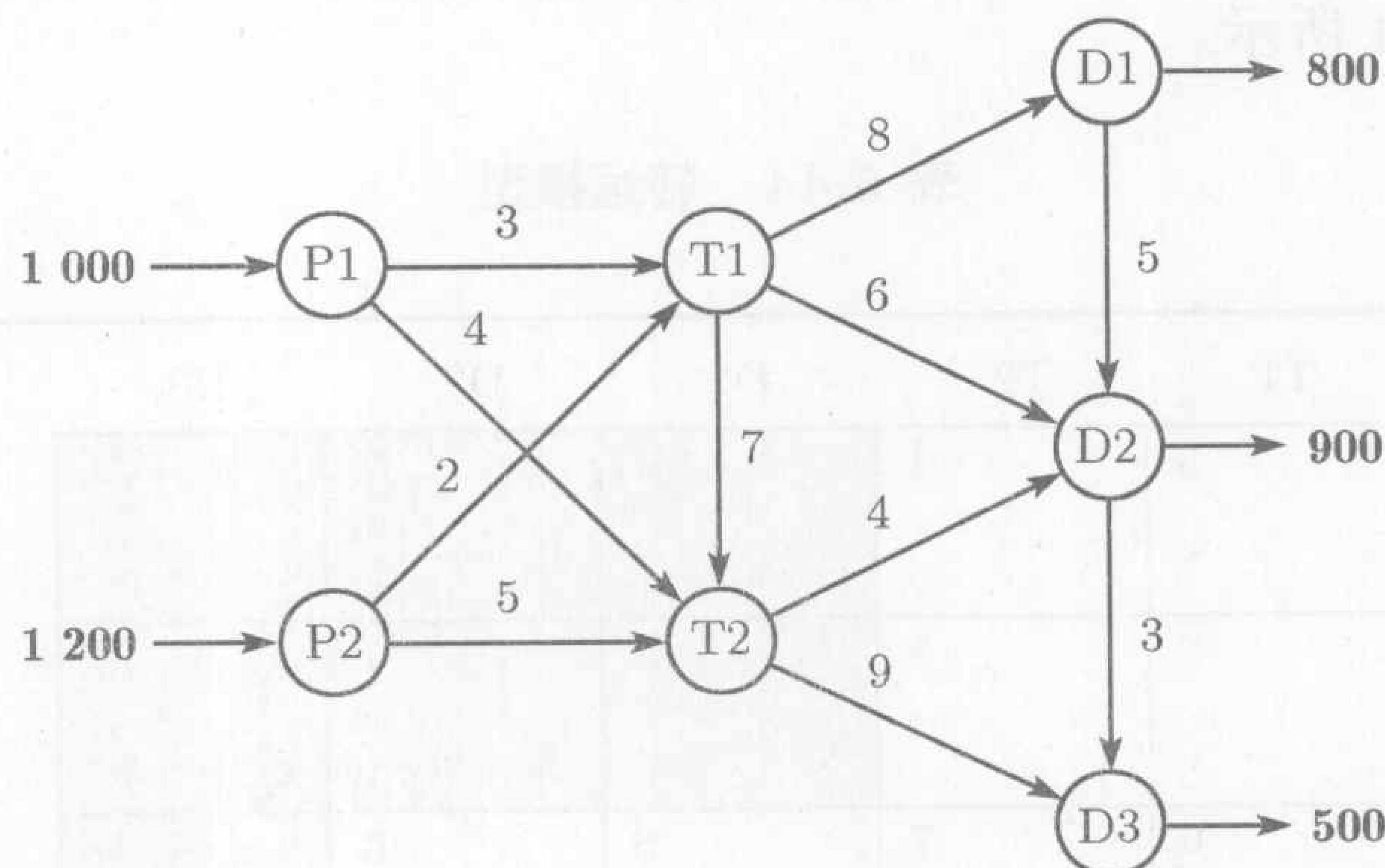


图 5.7 工厂与销售商之间的运输网络

在图 5.7 所示的网络中，因为制造厂 P1 和 P2 的全部供应量 2 200(=1 000+1 200) 辆汽车在最终到达目标节点 D1，D2，D3 以前，可以通过任何节点进行转运. 这样，网络中同时带有输入弧线和输出弧线 (T1，T2，D1，D2) 的每个节点既是起点，又是终点，被称为**转运节点**(transshipment node). P1 和 P2 被称为**纯供应节点** (pure supply node)，D3 被称为**纯需求节点** (pure demand node).

这个转运模型可转换成一个常规的运输模型，模型有 6 个起点 (P1，P2，T1，T2，D1，D2) 和 5 个终点 (T1，T2，D1，D2，D3). 不同节点的供应量和需求量可下列各式计算：

$$\text{每个纯供应节点的供应量} = \text{初始供应量}$$

$$\text{每个纯需求节点的需求量} = \text{初始需求量}$$

$$\text{每个转运节点的供应量} = \text{初始供应量} + \text{缓冲量}$$

$$\text{每个转运节点的需求量} = \text{初始需求量} + \text{缓冲量}$$

缓冲量应该充分大, 允许所有的初始供应 (或需求) 量能够通过任何一个转运节点. 设 B 为所需要的缓冲量, 则

$$\begin{aligned} B &= \text{全部供应量 (或需求量)} \\ &= 1\ 000 + 1\ 200 (\text{或} 800 + 900 + 500) \\ &= 2\ 200 \text{辆汽车} \end{aligned}$$

运用缓冲量 B 以及网络中给定的单位运输费用, 就可以构造出等价的常规运输模型, 如表 5.44 所示.

表 5.44 转运模型

	T1	T2	D1	D2	D3	
P1	3	4	M	M	M	**1 000**
P2	2	5	M	M	M	**1 200**
T1	0	7	8	6	M	B
T2	M	0	M	4	9	B
D1	M	M	0	5	M	B
D2	M	M	M	0	3	B
	B	B	**800 +** B	**900 +** B	**500**	

所得到的运输模型的解 (用 TORA 计算) 如图 5.8 所示. 注意转运的效果: 经销商 D2 收到 1 400 辆汽车, 留下 900 辆汽车满足需求, 把余下的 500 辆汽车转给经销商 D3.

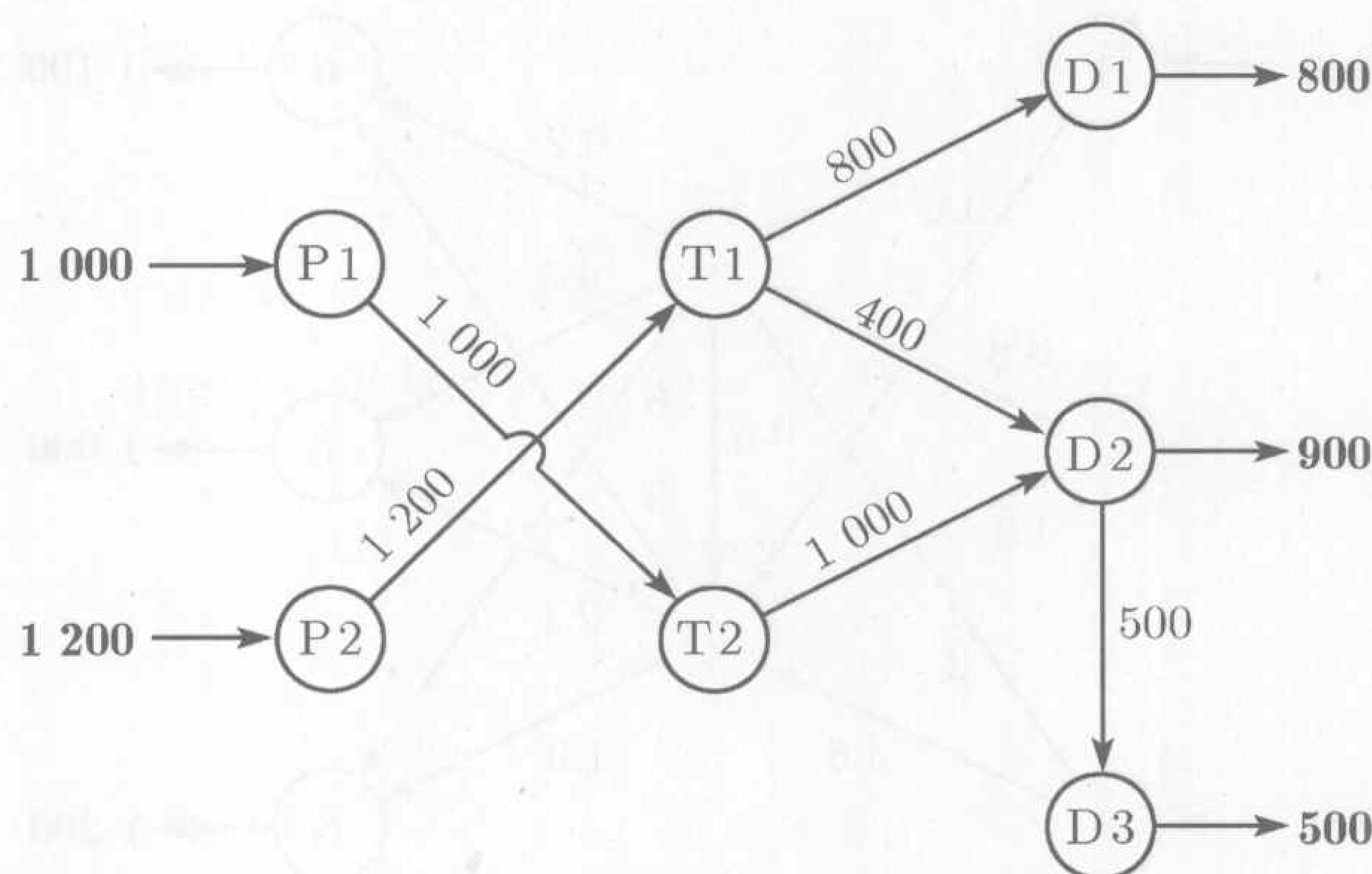


图 5.8 运输模型的解

习题 5.5A[①]

1. 图 5.9 的网络给出了从节点 1 和节点 2 经过节点 3 和节点 4 到节点 5 和节点 6 的运输路线. 单位运输费用表示在相应的弧上.

(a) 建立相应的转运模型.

(b) 求解本问题, 并表明从起点到终点的运输路线.

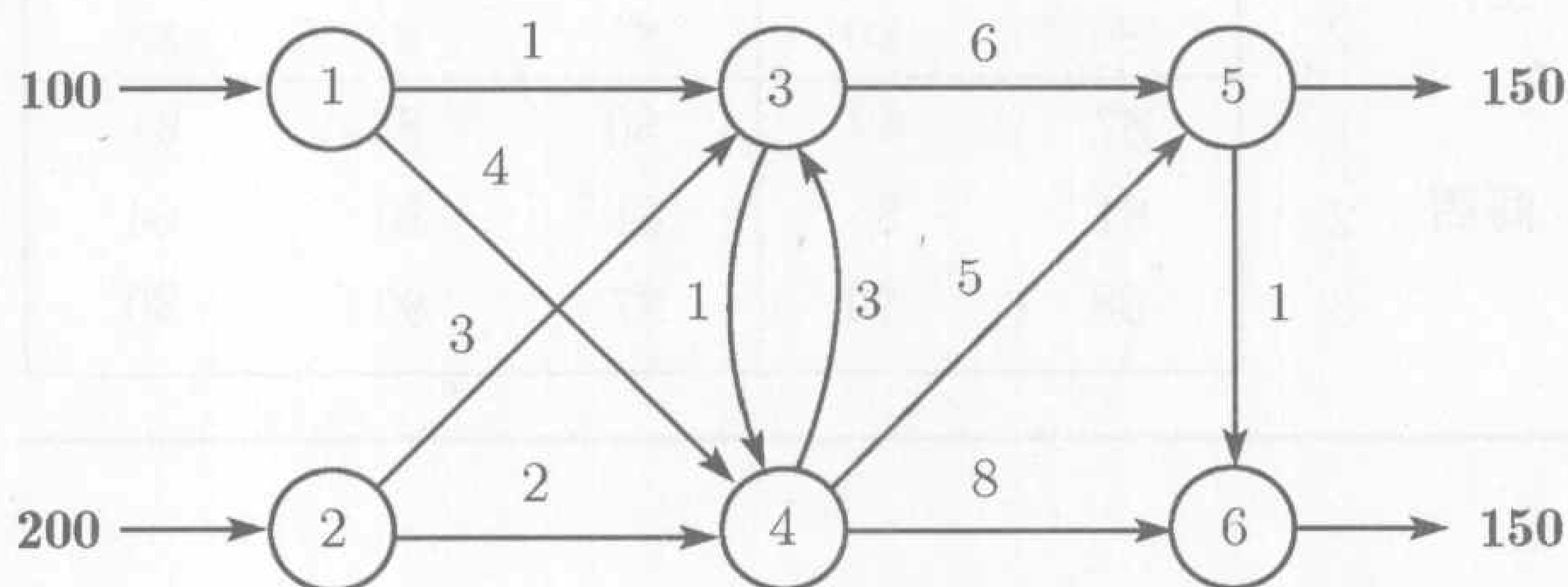


图 5.9 习题 5.5A 第 1 题的运输网络

2. 在第 1 题中, 假设起点 1 可连接到起点 2, 单位运输费用为$1, 并假定从节点 1 到节点 3 的单位运输费用增加到$5. 用转运模型描述本问题并求最优运输计划.

3. 图 5.10 所示的网络表明, 运输汽车的路线从 3 个工厂 (节点 1、节点 2 和节点 3) 出发, 通过分销中心 (节点 4 和节点 5) 送到 3 个经销商 (节点 6 到节点 8). 弧上标明了每辆汽车的运输费用 (以$100 为单位).

(a) 用转换模型解决此问题.

(b) 假设分销中心 4 能直接销售 240 辆汽车, 找到问题的最优解.

***4.** 考虑 2 个工厂为 3 个商店供应某种货物的运输问题. 起点 1 和起点 2 的可供应量为 200 件和 300 件, 商店 1, 2, 3 的需求量分别为 100 件、200 件、50 件. 货物在到达最终商店之前可以在工厂之间或商店之间流动. 根据表 5.45 所示的单位费用, 找到最优运输方案.

① 建议使用 TORA、Excel 规划求解或 AMPL 求解本习题.

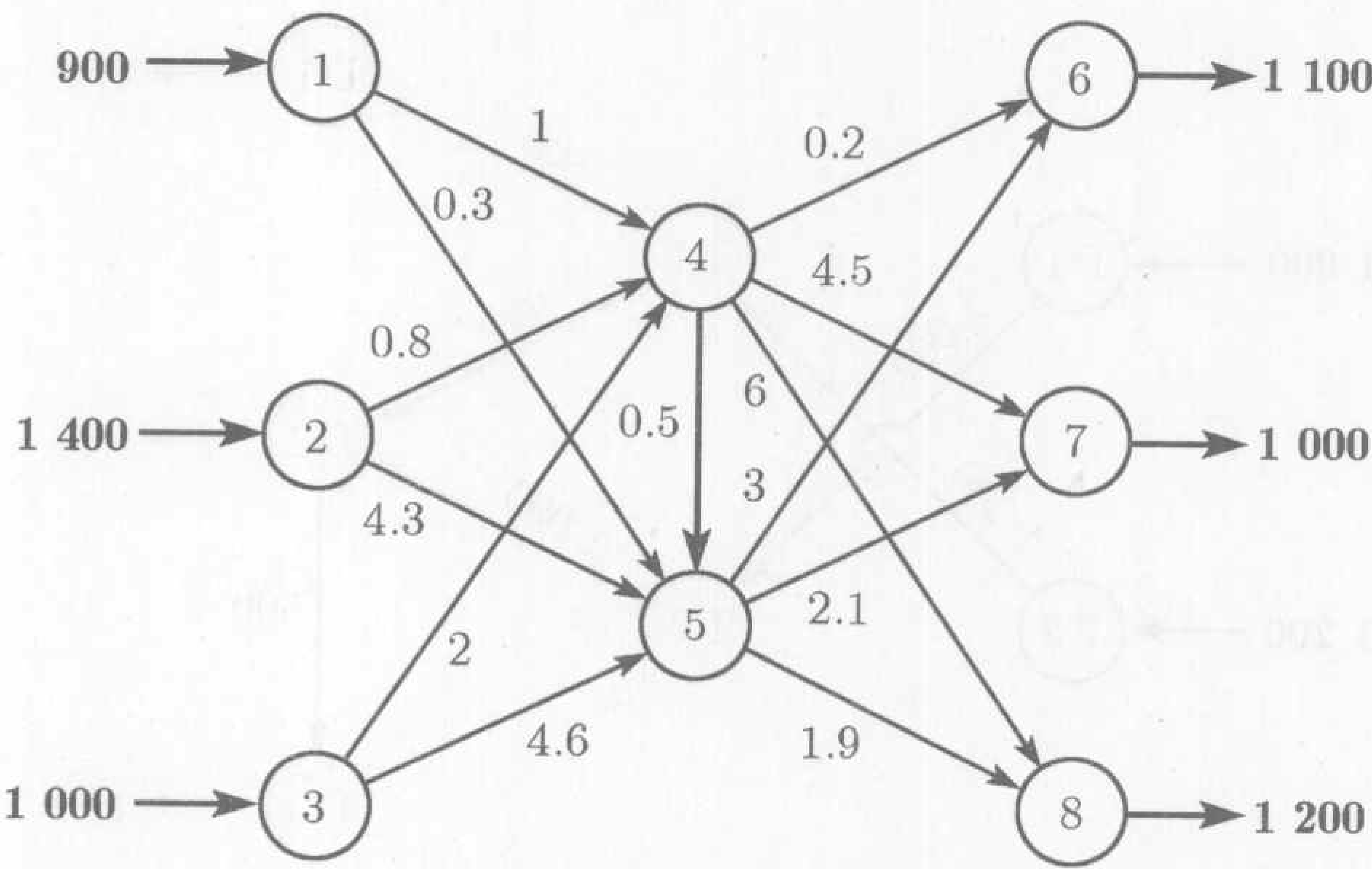


图 5.10 习题 5.5A 第 3 题的运输网络

表 5.45 第 4 题的数据

		工厂		商店		
		1	2	1	2	3
工厂	1	\$0	\$6	\$7	\$8	\$9
	2	\$6	\$0	\$5	\$4	\$3
商店	1	\$7	\$2	\$0	\$5	\$1
	2	\$1	\$5	\$1	\$0	\$4
	3	\$8	\$9	\$7	\$6	\$0

5. 考虑图 5.11 所示的输油管线网络. 不同的节点表示输送站和接收站. 网络中弧线上的数字表示了这些站点之间的距离, 两站之间每加仑石油的运输费用与管线长度成正比. 建立相应的转运模型, 并求最优解.

6. *最短路线问题*. 通过构造一个转运模型, 找到图 5.12 所示网络中节点 1 和节点 7 之间的最短路线. 不同节点之间的距离标示在网络上.(*提示*: 假定节点 1 有 1 个单位的净供应量, 节点 7 的净需求也是 1 个单位.)

7. 在例 5.5-1 的转运模型中, 定义 x_{ij} 为从节点 i 到节点 j 的运输量. 此问题可以构造成一个线性规划模型, 其中每个节点都产生一个约束方程. 建立这个线性规划模型, 并证明变量 x_{ij} 的约束系数 a_{ij} 具有以下特征:

$$a_{ij} = \begin{cases} 1, & \text{在约束 } i \text{ 中} \\ -1, & \text{在约束 } j \text{ 中} \\ 0, & \text{其他} \end{cases}$$

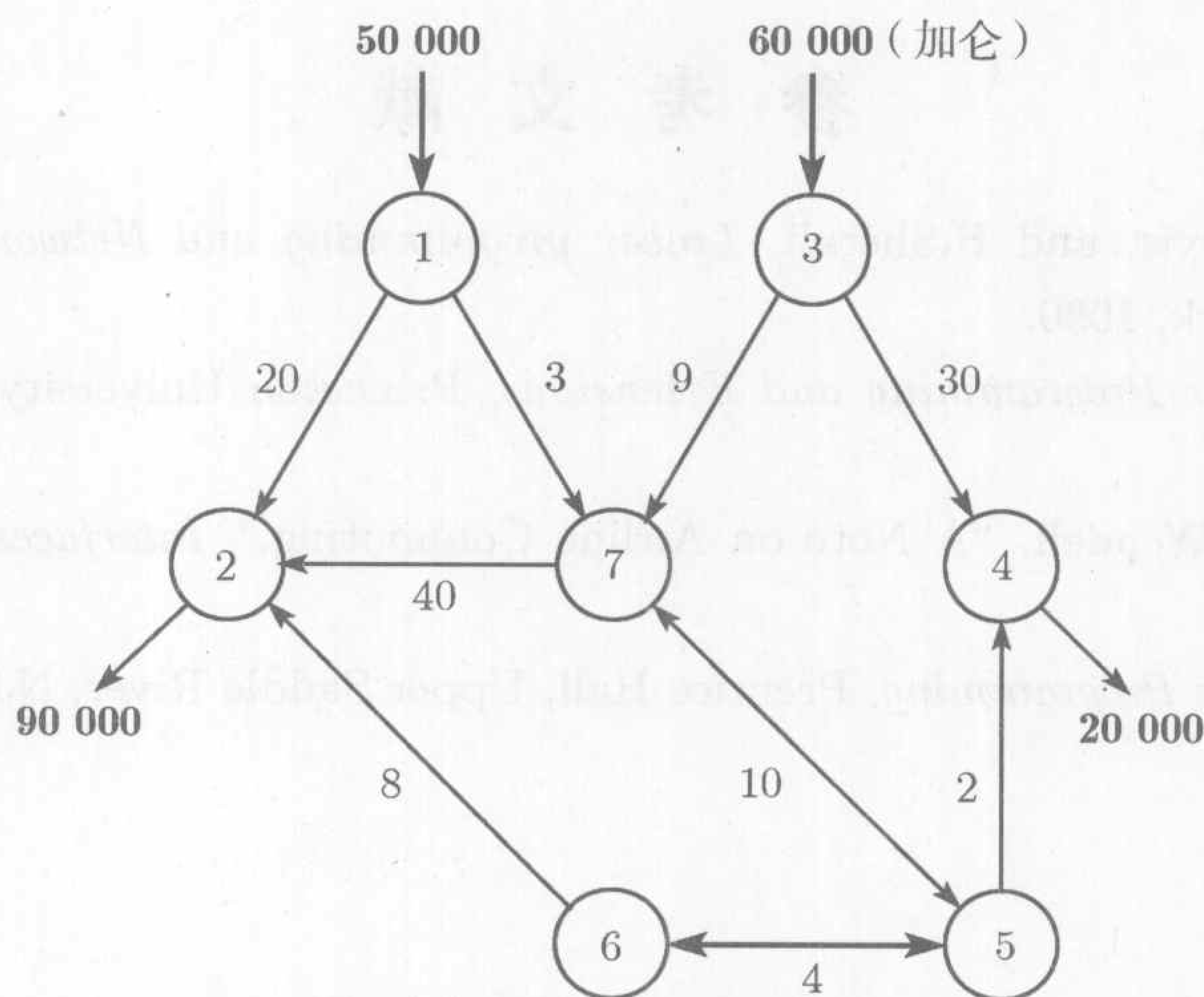


图 5.11 习题 5.5A 第 5 题的运输网络

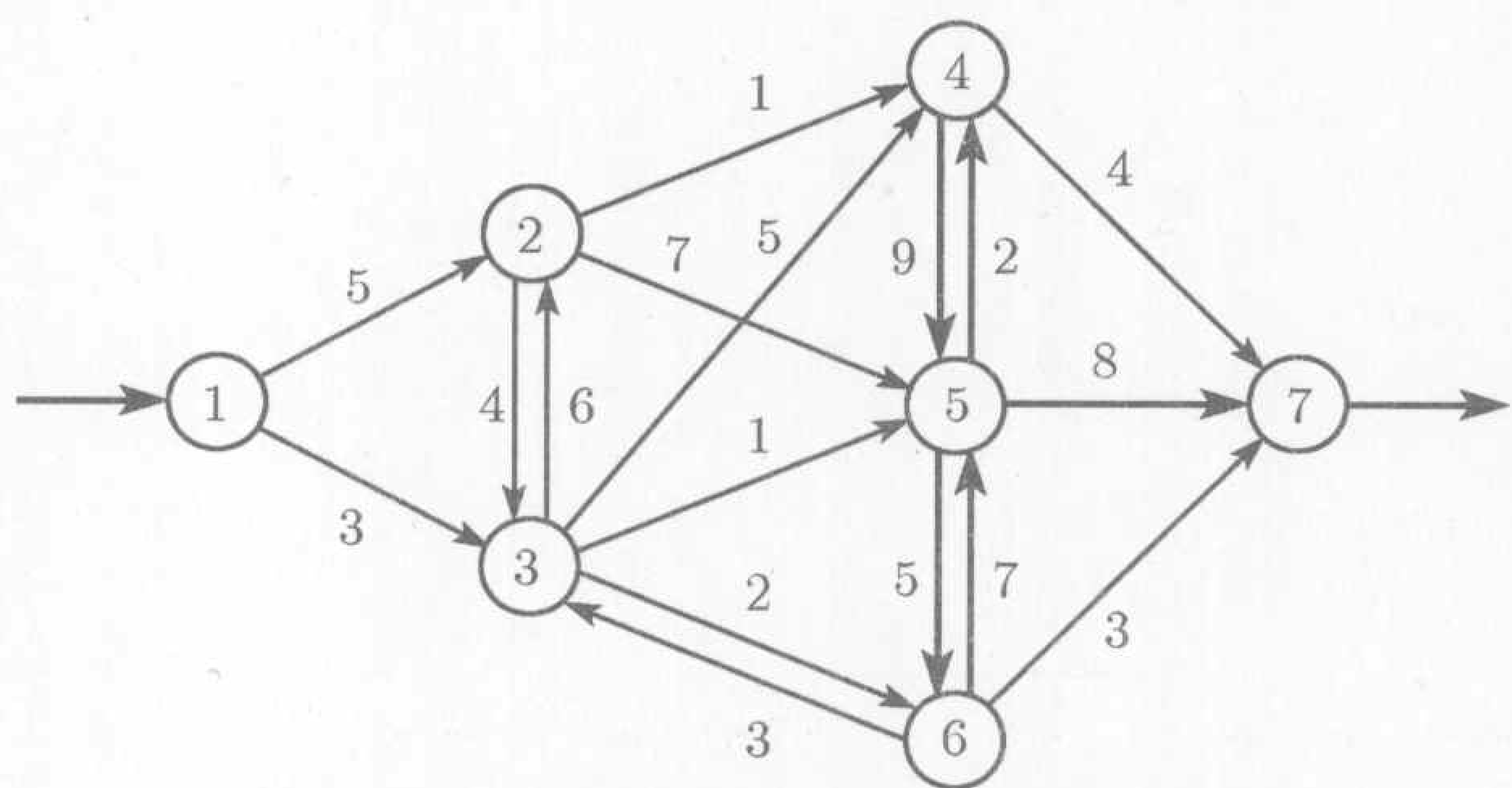


图 5.12 习题 5.5A 第 6 题的运输网络

8. 一个职业代理机构必须在以后的 5 个月里提供以下人数的劳工：

月份	1	2	3	4	5
劳工数	100	120	80	170	50

因为雇用劳工的费用与雇用时间长短有关, 在以后 5 个月的计划内如果能保留比需要数更多的劳工数可能更加经济. 下表估计了作为雇用时间长短函数的劳工费用 (美元)：

雇用月数	1	2	3	4	5
每位劳工的费用	100	130	180	220	250

把这个问题描述成一个线性规划, 然后, 对约束方程进行适当的代数处理, 证明这个模型可以转化为一个转运模型, 并找到最优解. (**提示**: 用第 7 题中的转运特点来把这个规划问题的约束转换为转运模型的约束)

参 考 文 献

Bazaraa, M.,J. Jarvis, and H.Sherali, *Linear programming and Network Flows*, 2nd ed., Wiley, New York, 1990.

Dantzig, G., *Linear Programming and Extensions*, Princeton University Press, Princeton, N.J., 1963.

Hansen, P. and R.Wendell, "A Note on Airline Commuting," *Interfaces*, Vol. 12, No. 1, pp.85-87, 1982.

Murty, K., *Network Programming*, Prentice Hall, Upper Saddle River, NJ, 1992.

第6章 网络模型

本章导读 本章所讲述的网络模型主要包括一些传统的应用，如寻找连接起一定数目地点的最优途径，寻找两个城市之间最短的路线，确定一个管道网络的最大流，确定一个网络中满足不同地点供应和需求的最小费用流，以及在一个项目中对各项活动进行规划的问题.

带有容量限制的最小费用流算法可以纳入到本章将要介绍的最短路径模型和最大流模型的广义网络，具体将在本书下册的 20.1 节中介绍.

在学习本章的时候，尤其要注意这些模型的非传统应用. 例如，最短路径模型可以用于求解最优的设备更换策略，最大流模型可以用于求解满足特定船只运输计划的最优船只数目问题，这些情形在本章中分别作为例题、习题或者案例加以介绍.

本章主要是利用线性规划建立网络模型并求解，建议读者认真理解它们之间的关系，因为大多数求解网络问题的商用软件都是基于线性规划的. 另外，对于一些需要增加边界约束的模型，也要求问题可以用线性规划求解.

为了对计算方面有进一步的了解，可以使用 TORA 交互式模块，本书中已介绍了它的求解过程步骤. 对于大型问题，本章也提供了不同算法设计的 Excel 规划求解和 AMPL 模型.

本章包括 1 个实际应用的综述、17 个例题、2 个 Excel 规划求解模型、3 个 AMPL 模型、69 个节后习题和 5 个案例. 所有的案例都放在本书下册的附录 E 中，AMPL/Excel/Solver/TORA 的程序放在文件夹 ch6Files 中.

实际应用——节约联邦政府的旅行费用

美国联邦政府的雇员需要到国内不同的地方参加发展会议和培训课程. 因为联邦雇员办公的地方分散在美国各个城市，所以会议主办城市的选择对于整个旅行预算有很大影响. 目前，选择会议和培训主办城市没有考虑旅行的花费. 相关的问题就是寻找确定主办城市的最优位置. 对于 1997 财政年，应用建立的模型至少节约了 40 万美元. 第 24 章中的案例 4 对这个问题有详细的研究.

6.1 网络模型的应用范围与定义

大量的运筹学问题都可以应用网络来建模求解 (节点之间通过边相连)：

(1) 在墨西哥湾上需要设计一个海面上的天然气管道网络，将各个井口连接到岸边的运输点. 模型的目标是最小化建设管道的费用.

(2) 在已经存在的道路网络中, 求两个城市之间的最短路径.

(3) 求从怀俄明的煤矿到休斯敦的发电厂的煤浆管道网络的最大运输量 (吨/年). (煤浆管道是通过一种特殊的管子将水注入到管道中来运输煤.)

(4) 给一个工程项目中的各项活动确定时间表 (开始日期和完成日期).

(5) 确定从油田到炼油厂的管道网络的最小费用流.

对于上面这些情形的求解均可以使用网络上的一些优化算法, 本章中将介绍其中的 4 个算法.

(1) 最小生成树算法 (第 1 种情形); (2) 最短路径算法 (第 2 种情形);

(3) 最大流算法 (第 3 种情形); (4) 关键路径算法 (第 4 种情形).

第 5 种情形对应的是网络上带有流量限制的最小费用流算法, 将在本书下册的 20.1 节中介绍.

网络模型的定义 一个网络是由**节点** (node) 集合以及连接节点的**弧**[arc, 或者边 (branch)] 集合组成. 用符号 (N, A) 表示, 其中 N 表示节点的集合, A 表示弧的集合. 例如, 图 6.1 中的网络可以描述为:

$$N = \{1, 2, 3, 4, 5\}$$
$$A = \{(1,2), (1,3), (2,3), (2,5), (3,4), (3,5), (4,2), (4,5)\}$$

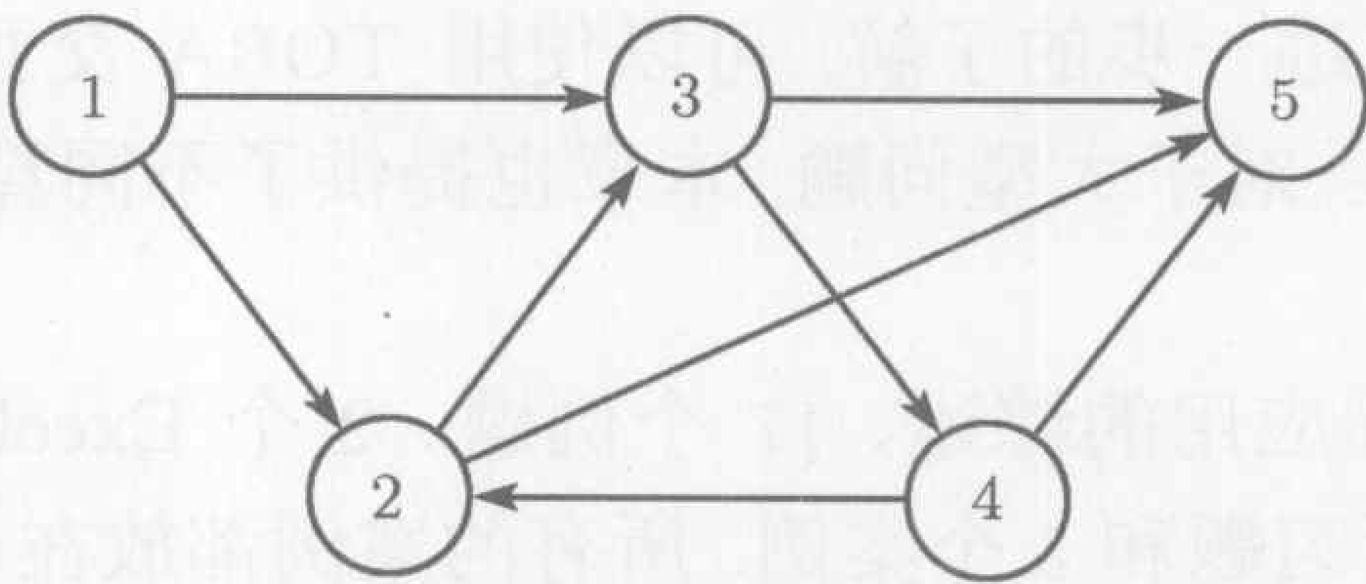


图 6.1 (N, A) 网络的例子

一个与网络相关的概念就是**流**(flow)(比如在管道中成品油的流量, 一条高速公路上车辆的流量). 一般来讲, 网络的流是由网络中弧上的容量限定的, 弧上的最大容量可以是有限的也可以是无限的.

如果一条弧上其中一个方向的流量允许是正数, 相反方向的流量为零, 那么这条弧称作是**有向的** (directed 或 oriented). 如果网络中所有的弧都是有向弧, 那么这个网络称作**有向网络** (directed network).

用一系列不考虑方向的弧将两个节点连接起来, 其中经过其他一些节点, 这一系列的弧就称作一条**路** (path). 如果一条路从某个节点出发经过其他一些节点又回到自身, 那么这条路称作**圈** (cycle) 或者**环** (loop). 例如, 图 6.1 中弧 $(2,3), (3,4), (4,2)$ 组成一个圈.

如果一个网络中任意两个不同的节点都至少有一条路连接, 那么这个网络称为**连通网络** (connected network). 图 6.1 给出的网络就是一个连通网络. 网络中部分

节点组成的一个无圈连通网络称作**树** (tree), 如果一棵树上包含了网络中所有的节点, 那么这棵树称作**生成树** (spanning tree). 图 6.2 给出了图 6.1 网络的一棵树和生成树的例子.

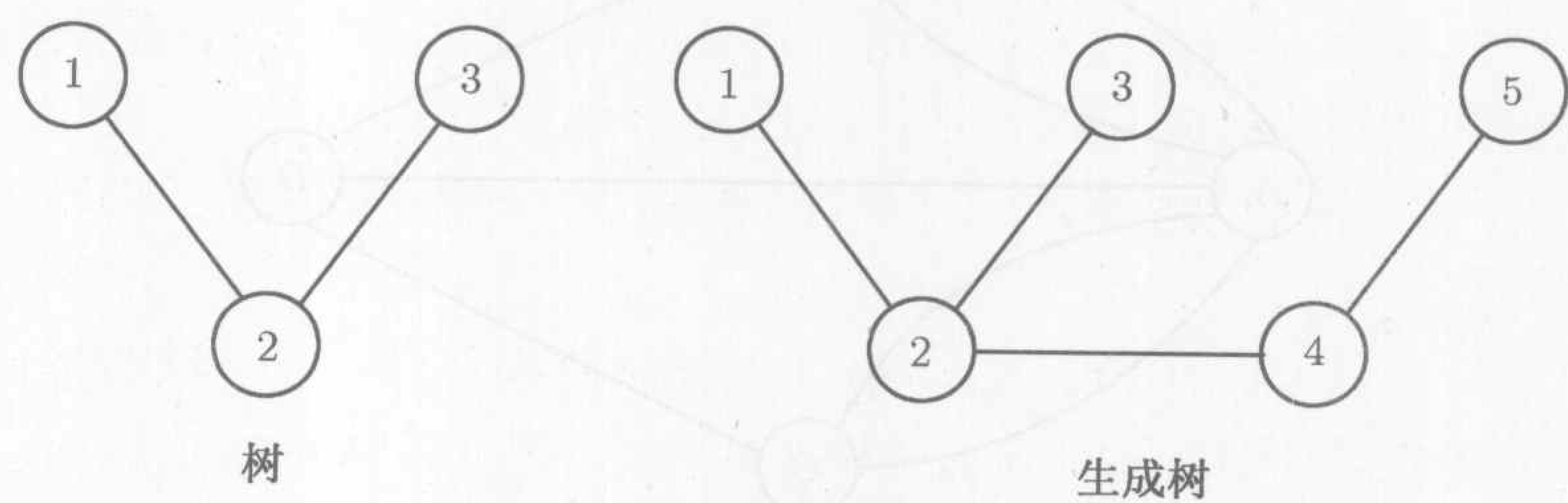


图 6.2 树和生成树的例子

例 6.1-1 (柯尼斯堡七桥问题)

普鲁士的柯尼斯堡 (现在的俄罗斯加里宁格勒城) 始建于 1254 年, 坐落在 Pergel 河两岸, 整个城市通过 7 座桥连接它的 4 个区域 (标号为 A, B, C, D), 如图 6.3. 该城市居民中一直流传这样一个问题: 是否可以找到一条到达每个区域并且通过每座桥恰好一次的环游(round trip). 每个区域上的到达次数不受限制.

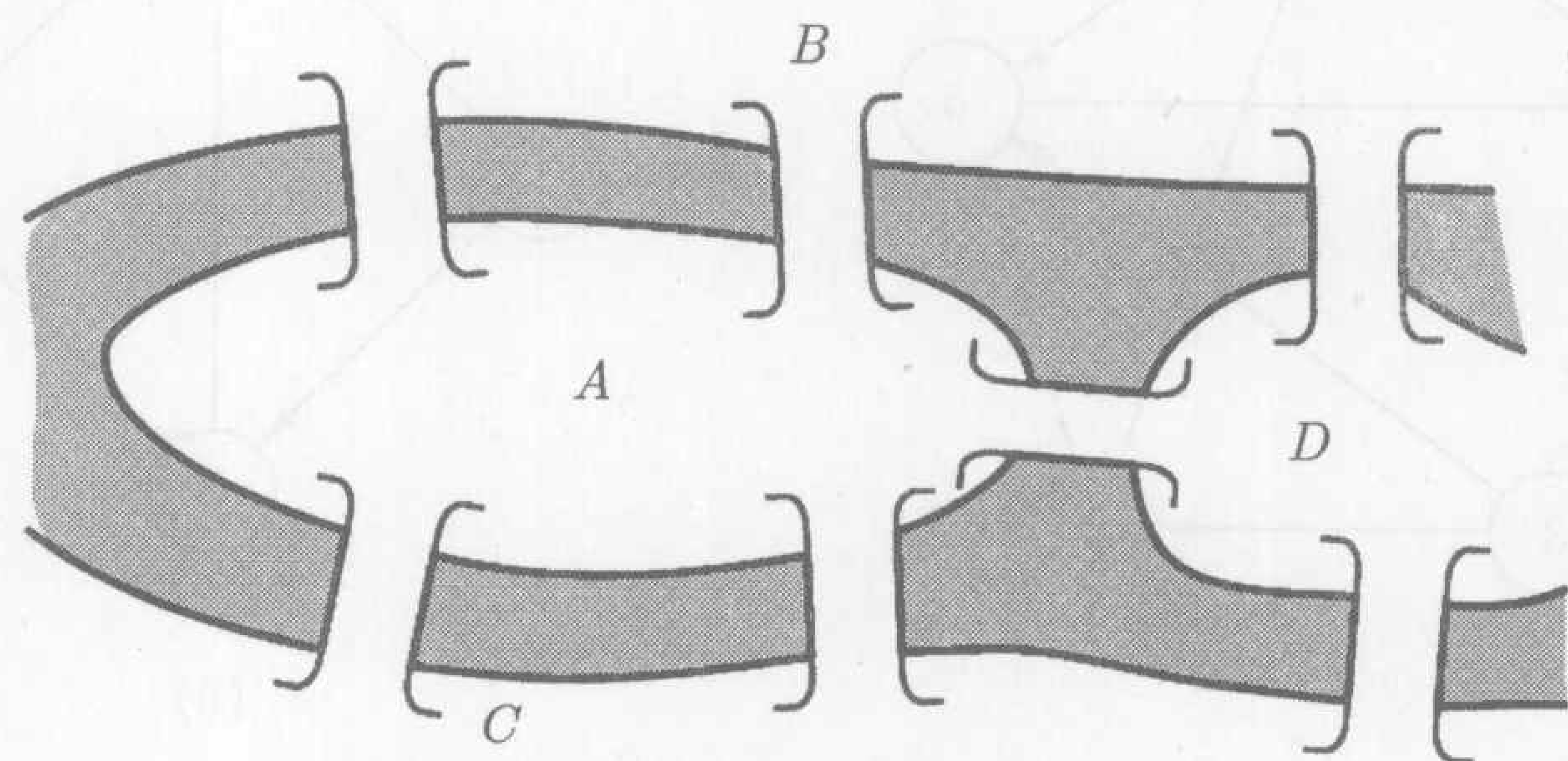


图 6.3 柯尼斯堡七桥问题

在 18 世纪中叶, 著名的数学家欧拉通过一种特殊的"路构造"方法证明了上面这个问题是不可能的. 到了 19 世纪初, 通过将这个问题转化为一个网络证明了同样的结论, 其中每个区域看作一个节点, 每座桥看作一条弧, 如图 6.4.

用基于网络的方法可以证明所设想的环游 (开始和结束于同一个区域) 是不存在的. 图中有 4 个节点, 每一个节点都有奇数条弧相连. 因为要求每座桥只能通过一次, 也就是图上的每条边只能通过一次, 但对于一个节点而言, 由于弧的数目是奇数, 所以不能保证进入和出去这个节点的次数相等①. 通过这个例题, 可以看出使

① 更一般的结论是: 一个网络中只有当每个节点都有偶数条边相连, 或者恰好只有两个节点有奇数条边相连, 才存在经过每条边恰好一次的通路; 否则就不可能存在. 参看 B. Hopkins 和 R. Wilson, "The Truth about Königsberg", *College Math Journal*, Vol.35, No. 3, pp. 198–207, 2004.

用网络表示能够很方便地求解问题.

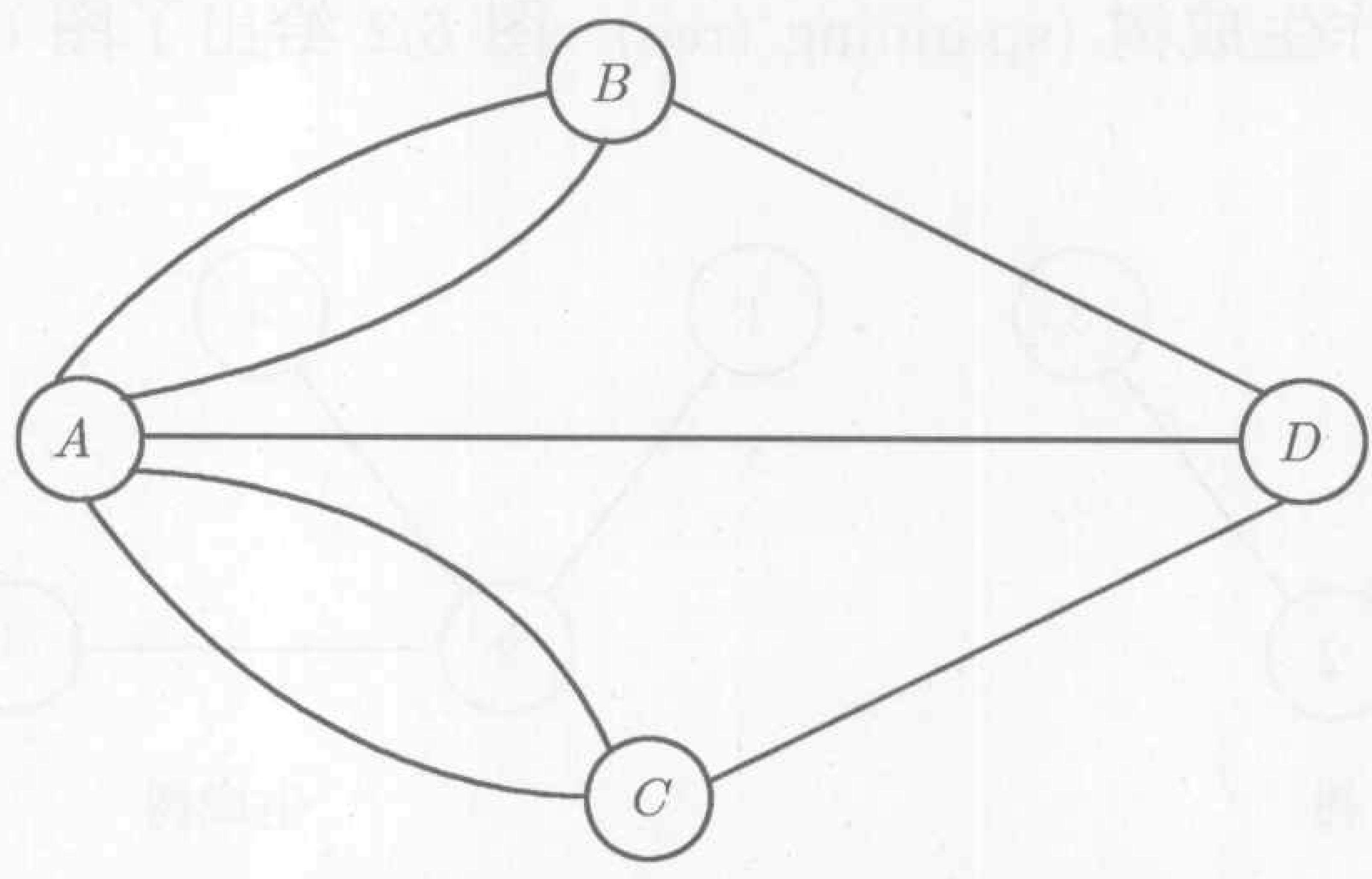


图 6.4 柯尼斯堡七桥问题的网络表示

习题 6.1A

*1. 指出图 6.5 中每个网络中的(a) 一条路; (b) 一个圈; (c) 一棵树; (d) 一棵生成树.

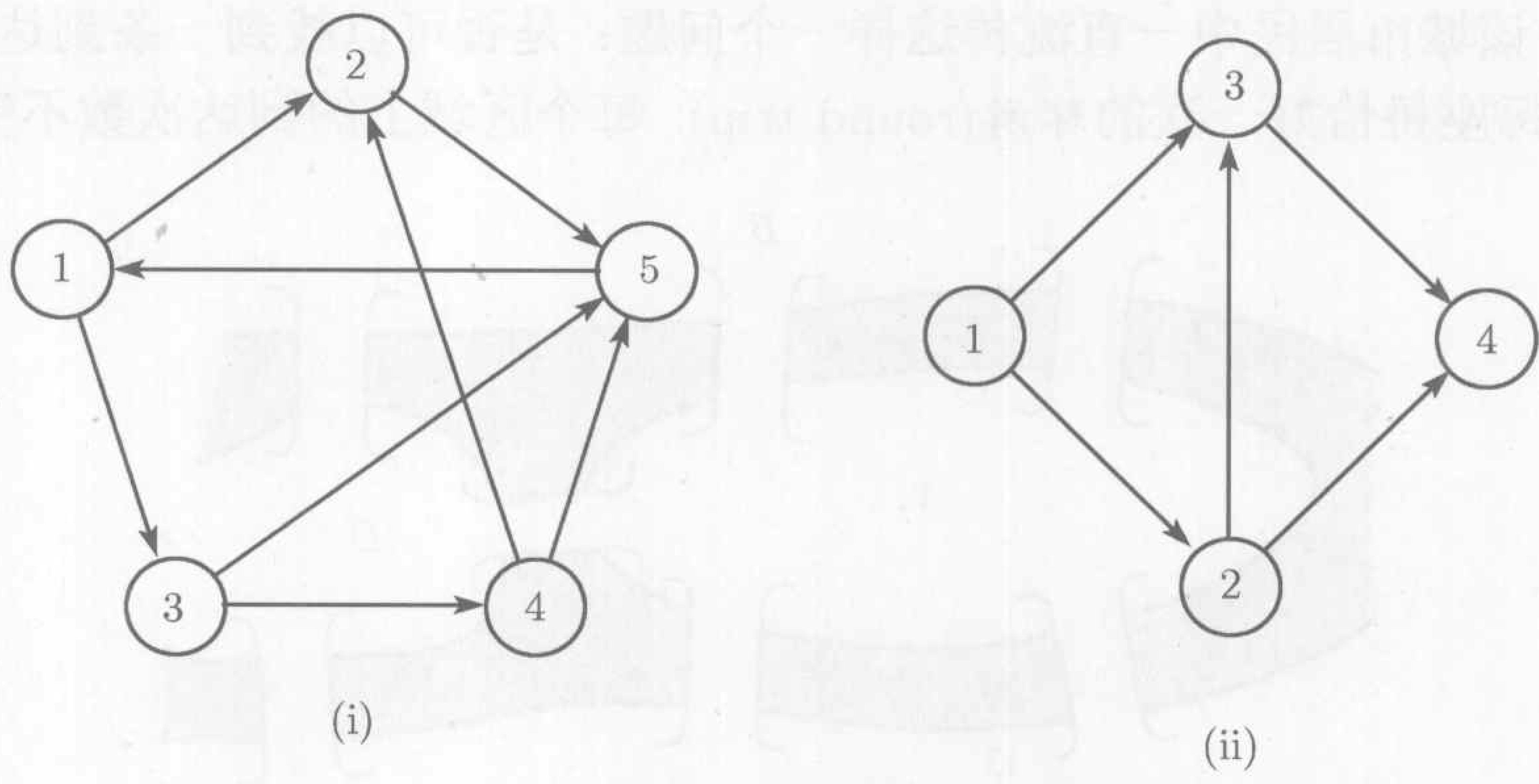


图 6.5 习题 6.1A 第 1 题中的网络

2. 指出图 6.5 中每个网络的节点集合 N 和弧集合 A.

3. 按下面给出的节点集合 N 和弧集合 A, 画出相应的网络.

$$N = \{1, 2, 3, 4, 5, 6\}$$
$$A = \{(1,2), (1,5), (2,3), (2,4), (3,4), (3,5), (4,3), (4,6), (5,2), (5,6)\}$$

*4. 将 8 个相同的正方形分别放成 3 行, 第 1 行 2 个, 第 2 行 4 个, 第 3 行 2 个, 使得每一行的正方形都是关于同一条竖直轴对称的. 需要将 1 ~ 8 这 8 个自然数分别填到上面 8 个正方形中, 使得任意水平、竖直、对角相邻的 2 个正方形中填上的数不相邻. 利用网络表示这个问题, 并用一个有步骤的方法求解.

5. 3 个警察押送 3 名犯人到一个小岛的监狱去服刑, 现在需要乘船渡过一条河, 这只船每次不能乘坐超过 2 人, 任何时候如果犯人的数目超过了警察的数目, 那么犯人就会造反. 利用网络模型设计一个乘船渡河的方案, 保证可以将犯人安全地押送到岛上.

6.2 最小生成树算法

最小生成树算法是用来连接一个网络的所有节点, 使树上边的总长度达到最小. 一个典型的应用实例就是, 需要在几个城镇之间修路, 使得任意两个城镇都有路相连, 中间可以穿过一个或者多个其他城镇, 那么最经济的道路设计方案就是使修路的里程数最小, 这个问题就可以应用最小生成树算法求解.

下面介绍具体的几个步骤. 令 $N=\{1,2,\cdots,n\}$ 是网络中节点的集合, 定义:

$$C_k = \text{在第 } k \text{ 步时已经连接起来的节点的集合}$$

$$\overline{C}_k = \text{在第 } k \text{ 步以后需要连接的节点的集合}$$

第 0 步 令 $C_0=\varnothing, \overline{C}_0=N$.

第 1 步 从 $\overline{C}_0$ 中的任意一个节点 i 开始, 令 $C_1=\{i\}$, 那么 $\overline{C}_1=N-\{i\}$. 设定 $k=2$.

一般的第 k 步 在还没有连接的节点集合 $\overline{C}_{k-1}$ 中选择一个节点 j^*, 使得 j^* 到 C_{k-1} 中某个节点之间的弧长最小. 然后将 j^* 放入 C_{k-1}, 从 $\overline{C}_{k-1}$ 中删除 j^*, 即

$$C_k = C_{k-1}+\{j^*\}, \quad \overline{C}_k=\overline{C}_{k-1}-\{j^*\}$$

如果未连接节点的集合 $\overline{C}_k$ 是空集, 那么停止. 否则, 设定 $k=k+1$, 重复这一步.

例 6.2-1

中西部有线电视公司正在计划为 5 个新的居民区提供有线电视服务, 图 6.6 给出了小区之间可以铺设电缆的情况以及相应的距离. 问题是确定最经济的电缆铺设方案, 使得 5 个小区可以连接起来.

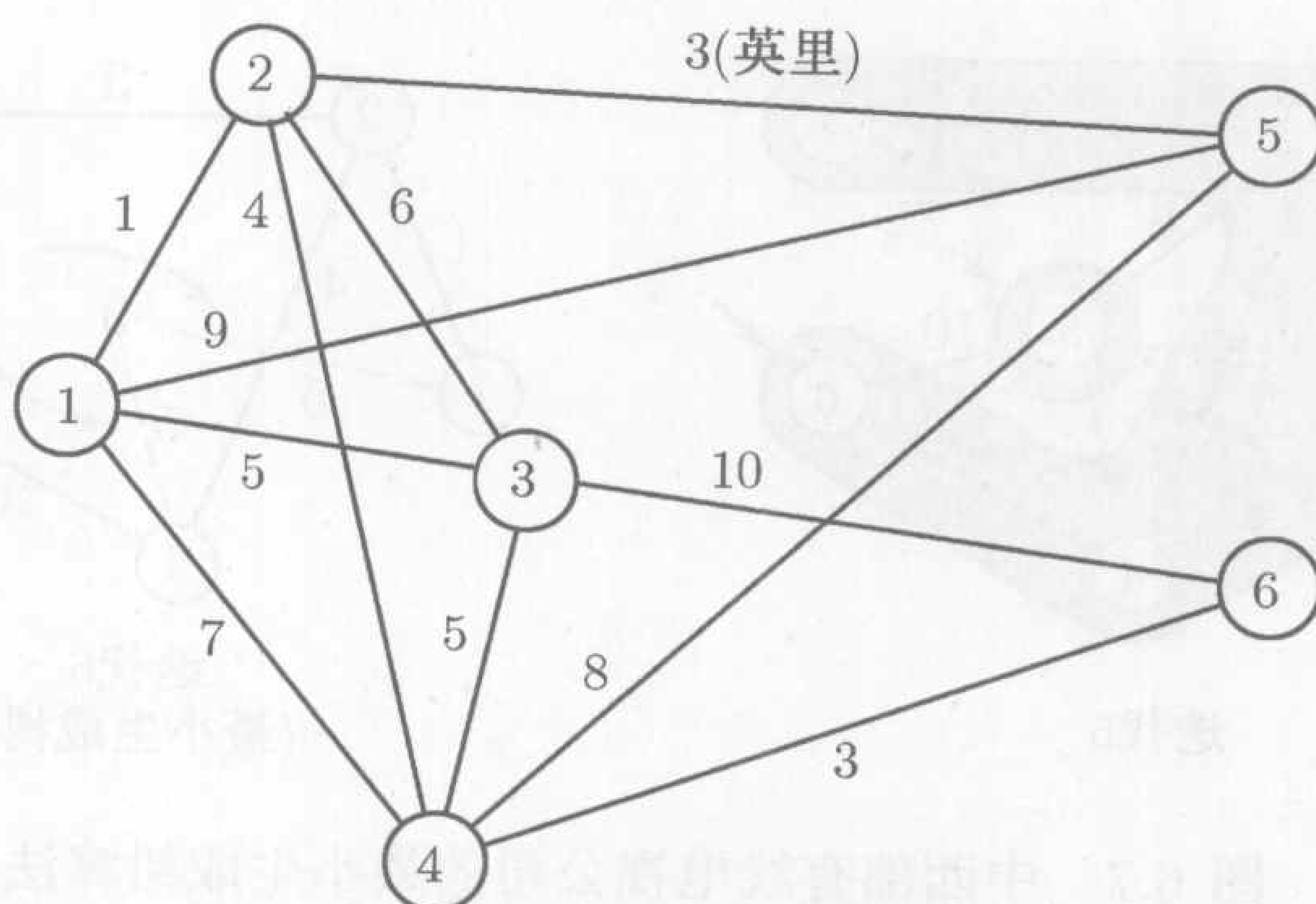


图 6.6 中西部有线电视公司的电缆连接网络

算法从节点 1 开始 (也可以选择其他任何一个节点), 那么

$$C_1 = \{1\}, \quad \overline{C}_1 = \{2, 3, 4, 5, 6\}$$

图 6.7 给出了算法的整个求解过程. 图中的细线表示所有连接集合 C 和集合 $\overline{C}$ 的候选边, 粗线表示已经选择作为连接集合 C 内节点的边, 虚线表示在当前一步选中的边. 例如, 第 1 步, 边 $(1, 2)$ 是从节点 1 到节点 2, 3, 4, 5, 6 所有候选边中最短的一条 (= 1 英里). 因此选取边 $(1, 2)$, $j^* = 2$, 那么,

$$C_2 = \{1, 2\}, \quad \overline{C}_2 = \{3, 4, 5, 6\}$$

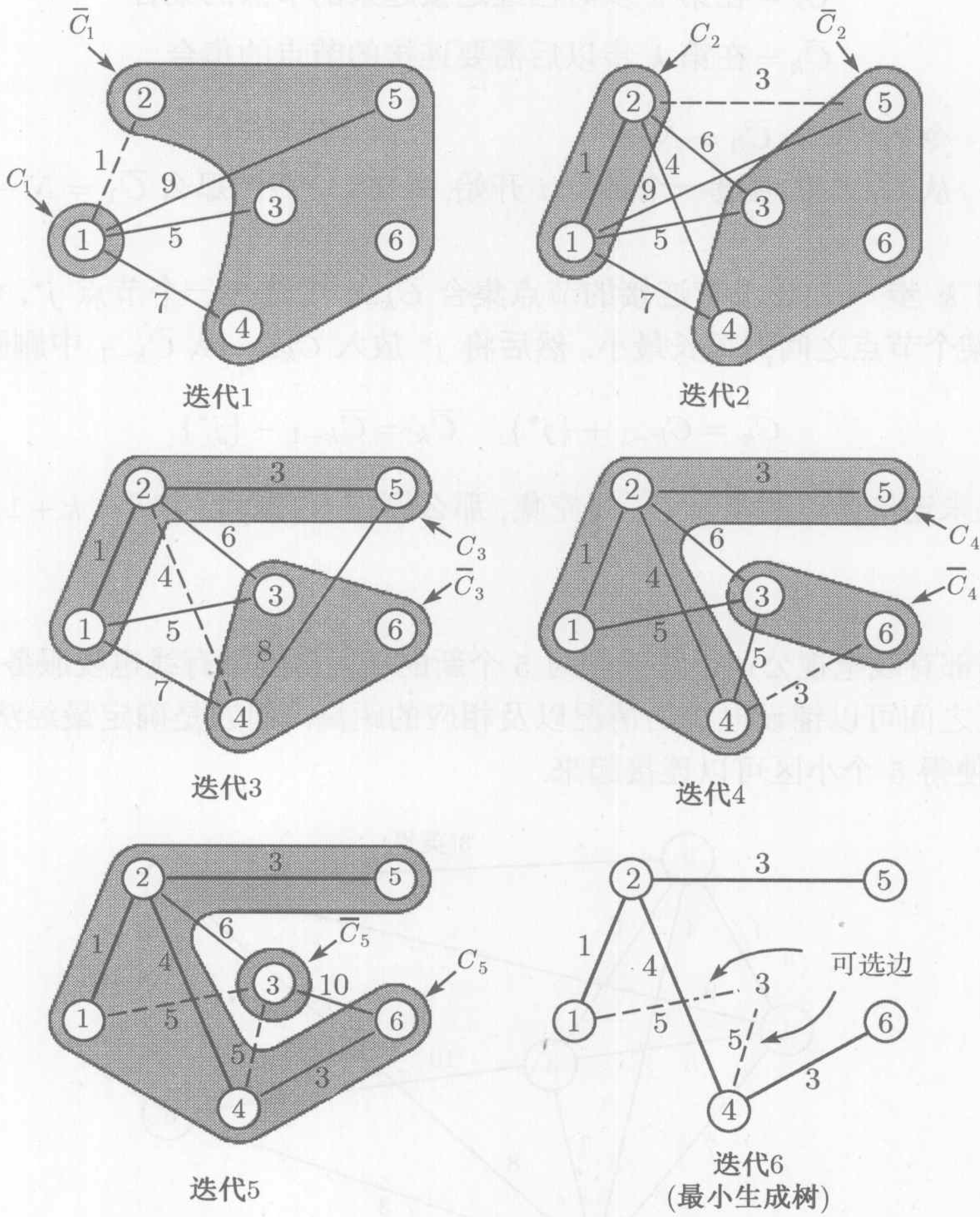


图 6.7　中西部有线电视公司的最小生成树算法

最小生成树算法在第 6 步就给出了问题的解, 如图 6.7. 要提供有线电视服务, 最少需要电缆的长度是 $1 + 3 + 4 + 3 + 5 = 16$.

TORA 程序

应用 TORA 可以得到最小生成树. 在 Mainmenu (主菜单) 中选择 Network models ⇒ Minimal spanning tree . 然后, 在菜单 SOLVER/MODIFY 中, 选择 Solve problem ⇒ Go to output screen . 在输出屏幕中选择一个 Starting node 再使用 Next iteration 或者 All iteration 来详细给出每一次迭代的情况, 当然可以选择其他新的 Starting Node 重新迭代. 文件 toraEx6.2-1.txt 中 TORA 使用的是例 6.2-1 的数据.

习题 6.2A

1. 以节点 5(代替节点 1) 作为初始节点求解例 6.2-1, 并说明算法得到的解是相同的.

2. 在下面的每个附加条件下求解例 6.2-1 的最小生成树:

*(a) 节点 5 与节点 6 之间可以通过长度为 2 英里的电缆连接.

(b) 节点 2 与节点 5 不能相连.

(c) 节点 2 与节点 6 之间可以通过长度为 4 英里的电缆连接.

(d) 节点 1 与节点 2 之间的需要的电缆长度变为 8 英里.

(e) 节点 3 与节点 5 之间可以通过长度为 2 英里的电缆连接.

(f) 节点 2 不能够与节点 3 和节点 5 直接相连.

3. 在美国的联合运输中, 装载货物的拖车需要使用一种特殊的火车在几个主要的城市之间的铁路上运送, 图 6.8 给出了美国几个主要的铁路枢纽, 以及现存的铁路路线. 现在要确定哪几条线路需要整修来满足枢纽之间的运送. 特别地, 洛杉矶 (LA) 需要直接与芝加哥 (CH) 相连, 来满足它们之间更繁重的运送需求. 其他的枢纽可以直接相连, 也可以间接相连. 目的是使得所选择的线路总长度最小 (单位: 英里). 然后指出在整修方案中包含哪几条线路.

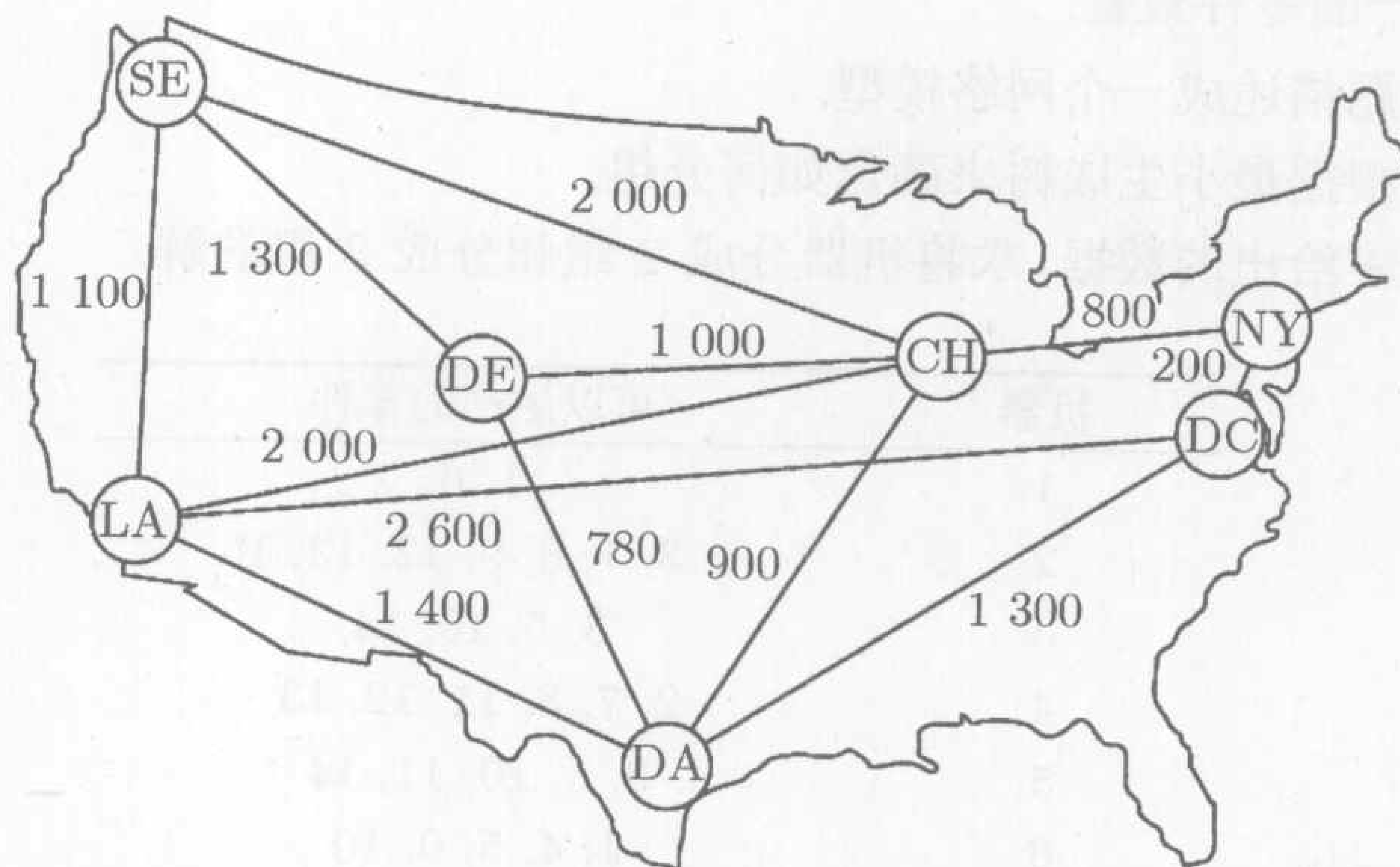


图 6.8 习题 6.2A 第 3 题对应的网络

4. 图 6.9 给出了 9 个海上天然气井口以及一个岸边运输点之间可以连接管道的情况. 由于井

口 1 距离岸边较近, 所以安装了一个容量足够大的泵来转输其他 8 个井口的天然气给运输点. 求连接 9 个井口到运输点的最小管道网络.

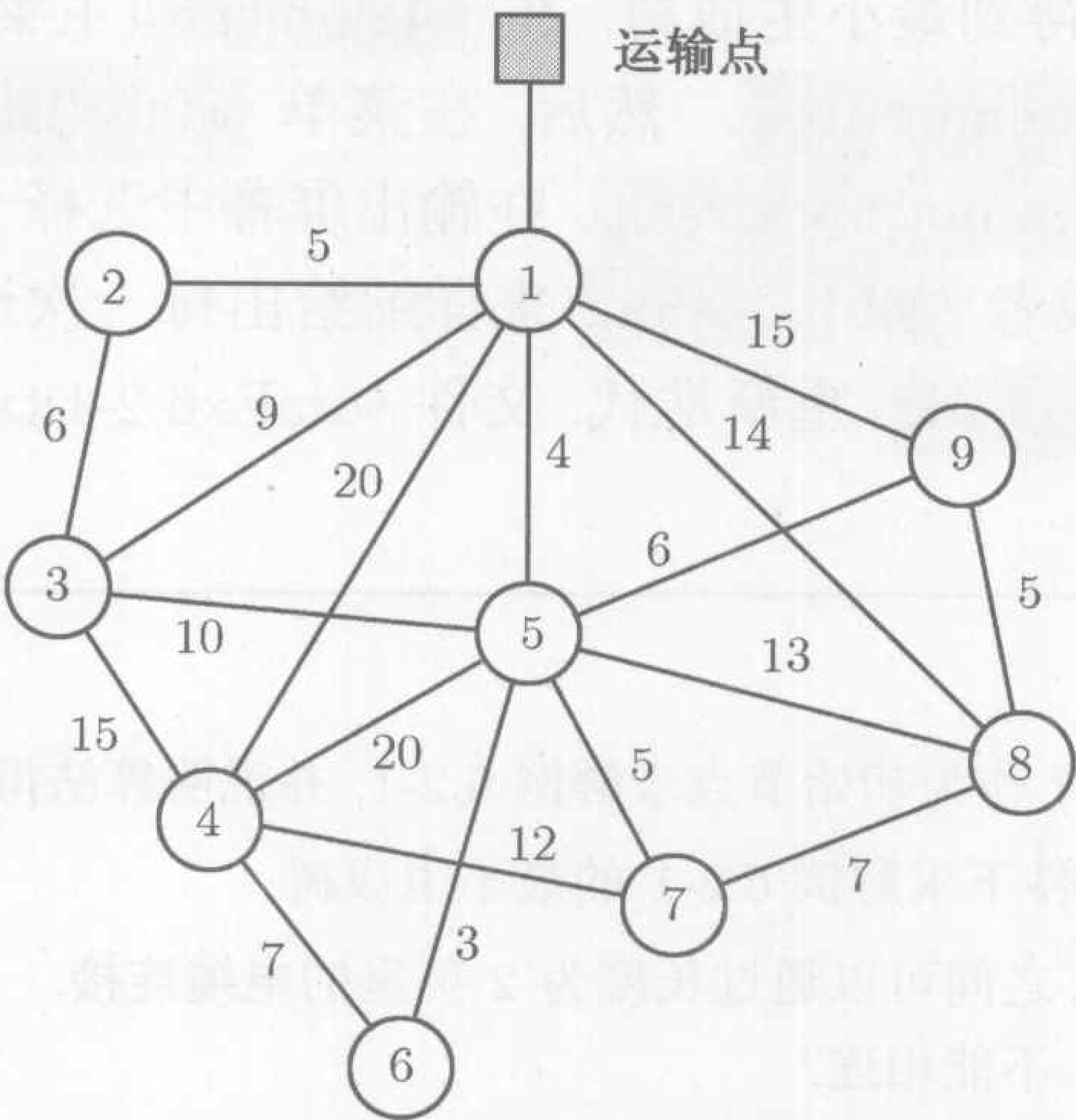


图 6.9 习题 6.2A 第 4 题对应的网络

***5.** 在第 4 题的图 6.9 中, 假定所有的井口根据天然气的压力可以分为两组: 高压组包括井口 2, 3, 4, 6, 低压组包括井口 5, 7, 8, 9. 由于压力不同, 不能将不同组的井口连接在一起. 同时, 两个组都需要通过井口 1 连接到运输点. 求在这种情形下的最小管道网络.

6. 某电子厂需要在 10 台机器上生产 15 种电子零件. 公司准备将机器分成若干个组, 使得零件在不同组上生产的"差异"最小, 零件在机器 i 和机器 j 上生产的"差异" d_{ij} 定义为

$$d_{ij} = 1 - \frac{n_{ij}}{n_{ij} + m_{ij}}$$

其中 n_{ij} 表示机器 i 和机器 j 上都可以生产的零件数量, m_{ij} 表示只能在机器 i 和机器 j 其中之一上生产的零件数量.

(a) 将这个问题描述成一个网络模型.

(b) 说明可以根据最小生成树来确定如何分组.

(c) 根据下表中给出的数据, 求将机器分成 2 组和分成 3 组的解.

机器	可以生产的零件
1	1, 6
2	2, 3, 7, 8, 9, 12, 13, 15
3	3, 5, 10, 14
4	2, 7, 8, 11, 12, 13
5	3, 5, 10, 11, 14
6	1, 4, 5, 9, 10
7	2, 5, 7, 8, 9, 10
8	3, 4, 15
9	4, 10
10	3, 8, 10, 14, 15

6.3 最短路径问题

最短路径问题是在一个运输网络中求一条从源点到目的点的最短路径, 通过下面的例题, 可以看出许多其他的问题也可以转化成这个模型.

6.3.1 最短路径应用的实例

例 6.3-1 (设备更换问题)

RentCar 公司开展了一项 4 年一周期的汽车更换政策. 在每一年的开始, 消费者都可以选择继续使用购买的汽车, 也可以选择更换汽车. 一辆汽车使用最少 1 年, 最多 3 年, 下表给出了汽车更换的费用, 它取决于这辆车购买时的年数以及使用的年限.

开始购买汽车的年数	给定使用年限的更换费用 (美元)		
	1	*2*	*3*
1	4 000	5 400	9 800
2	4 300	6 200	8 700
3	4 800	7 100	—
4	4 900	—	—

这个问题可以描述成一个网络模型, 节点 1 到节点 5 表示第 1 年到第 5 年每年的开始. 与节点 1 (第 1 年) 有弧相连的节点只有 2, 3, 4, 因为按规定汽车的使用年限为 1 到 3 年, 同理, 可以推断出从其他节点出发的弧. 每条弧的长度等于更换费用. 那么这个问题的解就等价于求一条从节点 1 到节点 5 的最短路径.

如图 6.10 给出了相应的网络. 利用 TORA①, 可以得到最短路径 (途中用粗线表示) 是 $1 \to 3 \to 5$. 这个解说明了汽车需要在第 1 年 (节点 1) 购买, 然后使用 2 年, 在第 3 年 (节点 3) 开始的时候更换新的汽车, 更换得到的汽车一直使用到第 4 年底. 那么这个更换策略总的花费是 12 500 美元 (= \$5 400 + \$7 100).

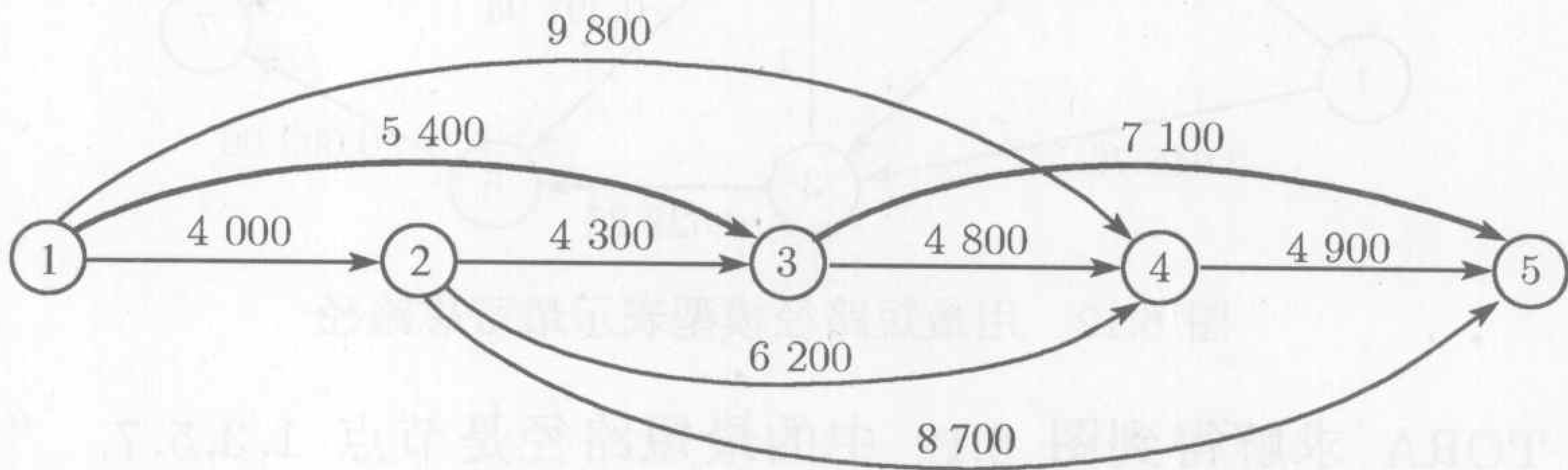


图 6.10 设备更换问题的最短路径模型

① 从 Main menu 选择 Network models ⇒ Shortest route. 从 SOLVE/MODIFY 菜单中选择 Solve problem ⇒ Shortest routes.

例 6.3-2 (最可靠路径)

I. Q. Smart 每天开车去工作时, 他能够通过网络分析找到最短的行程. 但是, 在这样的路上行驶往往会被限速, 如果超速的话, 就会被巡逻的警察罚款, 所以最短的路径不一定是最好的选择. Smart 因此会选择走一条路径, 它不被警察扣罚的概率为最大.

图 6.11 中的网络给出了住址和工作地点之间的可能路线, 并给出与各路段相应的不被扣罚的概率. 一条路径不被扣罚的概率是这条路径上每一段上不被扣罚的概率的乘积. 例如, 对于路径 $1 \to 3 \to 5 \to 7$ 不会被扣罚的概率是 $0.9 \times 0.3 \times 0.25 = 0.067\,5$. Smart 的目标是选择一条不会被扣罚的概率最大的路径.

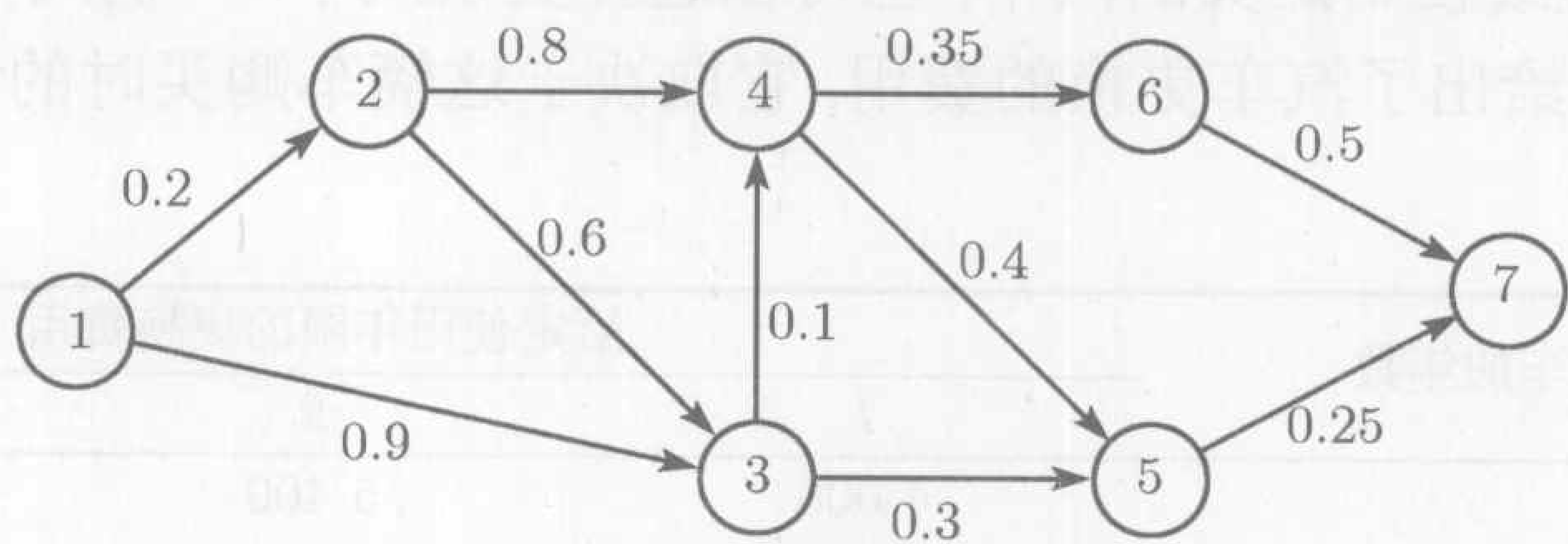


图 6.11　最可靠路径的网络模型

通过对数变换, 这个问题可以转化为最短路径问题, 也就是将概率乘积转化成概率对数的和, 例如不被扣罚的概率为 $p_{1k} = p_1 \times p_2 \times \cdots \times p_k$, 那么取对数之后是 $\lg p_{1k} = \lg p_1 + \lg p_2 + \cdots + \lg p_k$.

从数学的角度来看, p_{1k} 的最大化等价于 $\lg p_{1k}$ 的最大化. 因为 $\lg p_{1k} \leqslant 0$, 所以 $\lg p_{1k}$ 的最大化又等价于 $-\lg p_{1k}$ 的最小化. 根据这个变换, 对所有的 j, 图 6.11 中的每条边上的概率 p_j 可以用 $-\lg p_j$ 来代替, 这就产生了图 6.12 的最短路径网络问题.

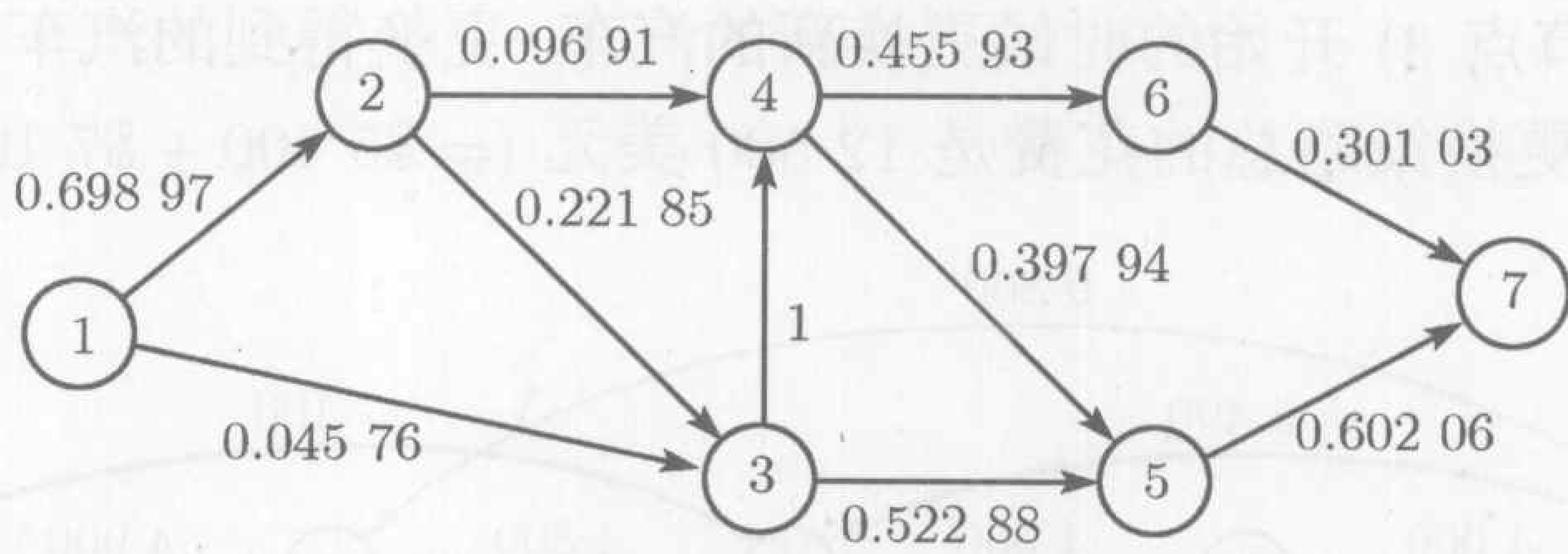


图 6.12　用最短路径模型表示最可靠路径

应用 TORA 求解得到图 6.12 中的最短路径是节点 1, 3, 5, 7, “长度”为 $1.170\,7(= -\lg p_{17})$. 那么对应这条路径的不被扣罚的概率只有 p_{17}=0.067 5, 或许对于 Smart 来说这也还不是一条理想的路径.

例 6.3-3 (3 把水壶问题)

现在有一把容量为 8 加仑的水壶装满了水, 借助于一把容量为 5 加仑和一把

容量为 3 加仑的空水壶将这 8 加仑的水平均分成两份. 不允许借助其他的工具, 如何能用最少次数的转移达到这个目的?

当然可以通过直观的方法猜测这个问题的解, 实际上, 这个问题可以使用最短路径模型有步骤地表示出来.

用一个节点表示 8 加仑、5 加仑和 3 加仑容量水壶中的水, 这意味着网络开始于节点 (8, 0, 0), 结束于我们想得到的节点 (4, 4, 0). 通过从某个水壶向另外一个水壶倒水, 我们可以从当前的节点移动到新的节点.

图 6.13 展示了从起点 $(8,0,0)$ 到终点 $(4,4,0)$ 的不同路径. 两相邻节点之间的弧代表一次转移, 因此可假设每条弧长为 1. 问题就化成了节点 $(8,0,0)$ 与节点 $(4,4,0)$ 之间的最短路径问题.

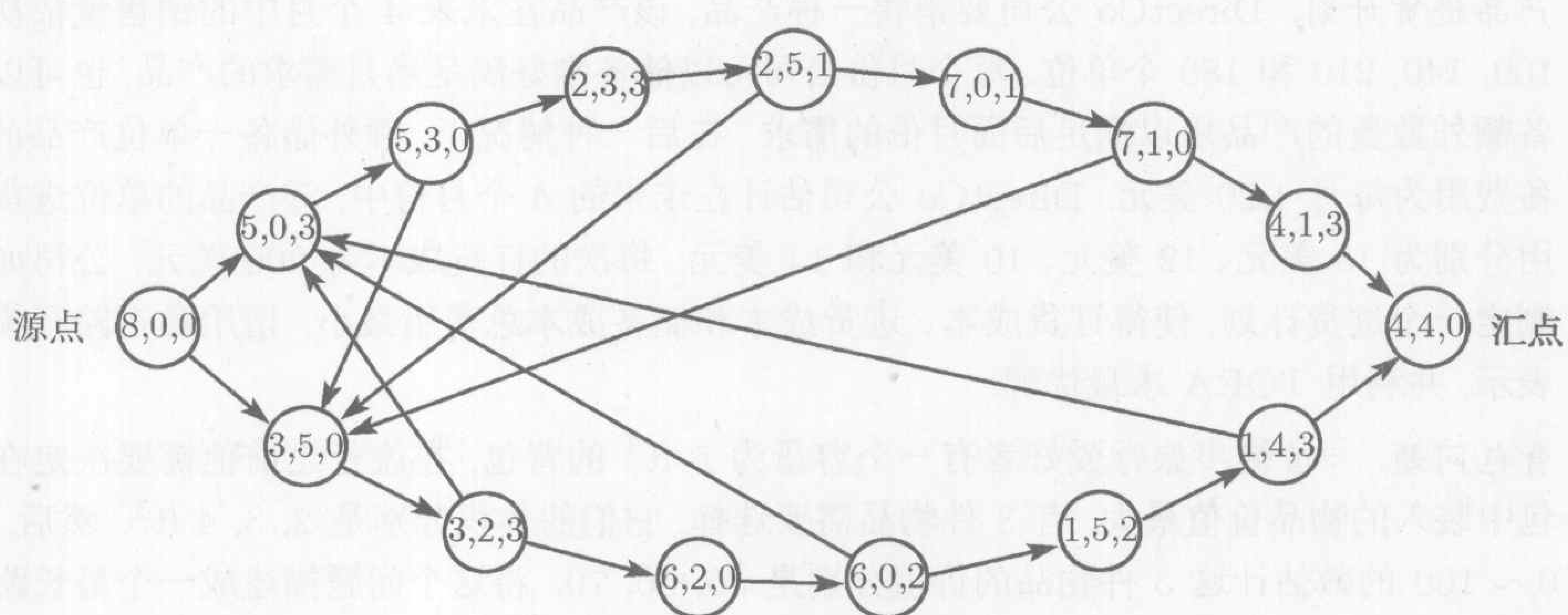


图 6.13 用最短路径模型表示 3 把水壶问题

问题的最优解就是图 6.13 给出的最下面的一条路径, 需要 7 次转移.

习题 6.3A

1. 按照下面的假设重新构造例 6.3-1 中的设备更换模型. 假定一辆汽车使用至少 2 年, 最多 4 年. 一个规划周期从第 1 年年初开始, 到第 5 年年底为止. 下表给出了必要的数据信息.

开始购买汽车的年数	给定使用年限的更换费用 (美元)		
	2	*3*	*4*
1	3 800	4 100	6 800
2	4 000	4 800	7 000
3	4 200	5 300	7 200
4	4 800	5 700	—
5	5 300	—	—

2. 图 6.14 是两个站点 1 和 7 之间的通信网络. 每条边上的数表示消息成功发送的概率. 我们从站点 1 向站点 7 发消息, 希望找到一条路径, 使得站点 7 成功接收到消息的概率最大. 请用最短路径模型描述, 并给出最优解.

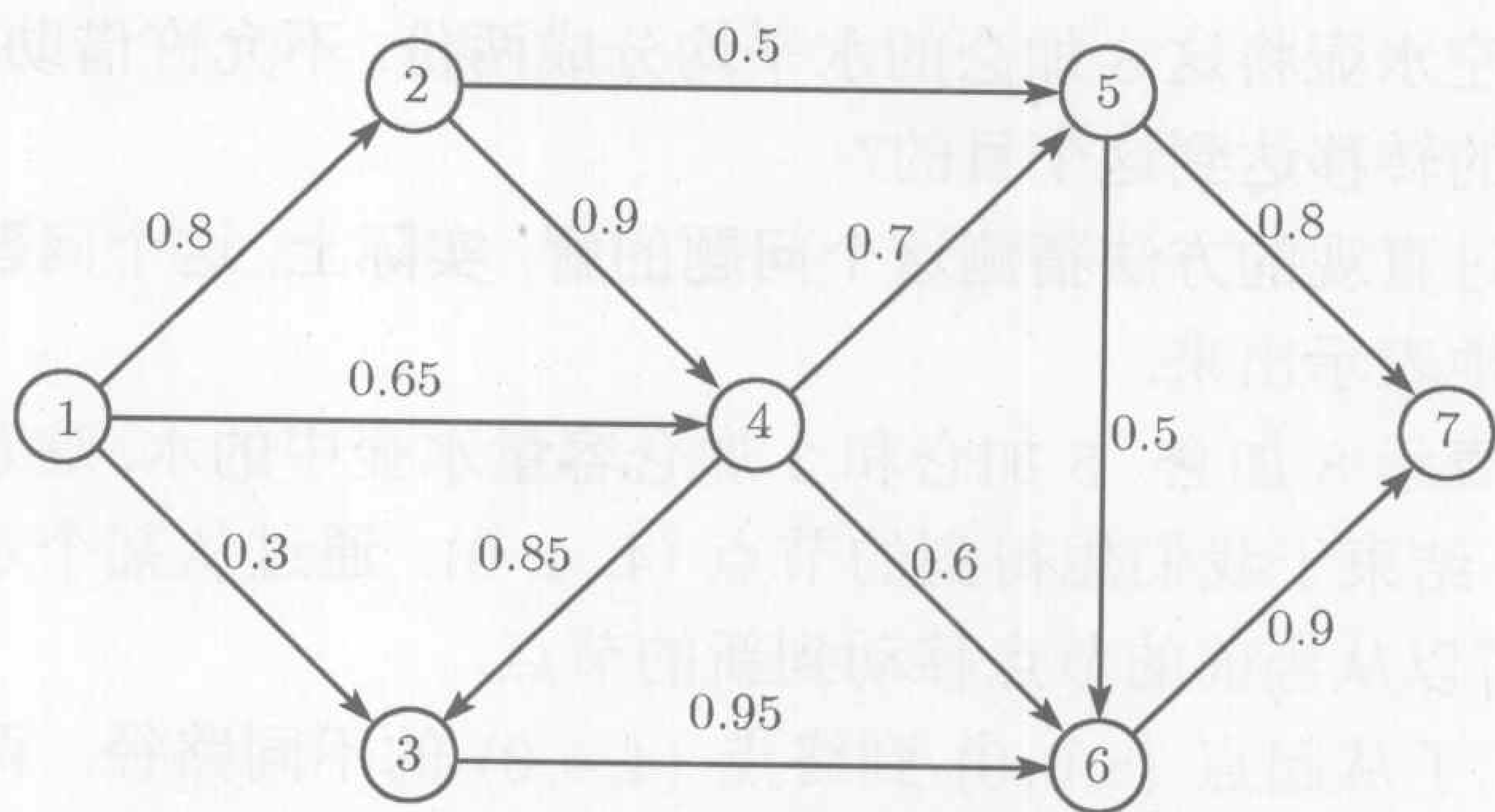


图 6.14　习题 6.3A 第 2 题对应的网络

3. *产品进货计划*. DirectCo 公司要销售一种产品, 该产品在未来 4 个月中的销售量依次为 100, 140, 210 和 180 个单位. 每个月份公司可以储备刚好满足当月需求的产品, 也可以储备额外数量的产品用以满足后面月份的需求. 在后一种情况下, 额外储备一单位产品的储备费用为每月 1.20 美元. DirectCo 公司估计在未来的 4 个月份中, 该产品的单位进货费用分别为 15 美元、12 美元、10 美元和 14 美元. 每次的订货成本为 200 美元. 公司如何制定一个进货计划, 使得订货成本、进货成本和储备成本总费用最小. 请用最短路径模型表示, 并利用 TORA 求最优解.

***4.** *背包问题*. 一位徒步旅行爱好者有一个容量为 5 ft^3 的背包, 在旅行之前他需要决定在背包中装入的物品价值最大, 有 3 件物品需要选择, 它们的体积分别是 2, 3, 4 ft^3. 然后, 用 $0 \sim 100$ 的数估计这 3 件物品的价值分别是 30, 50, 70. 将这个问题描述成一个最长路径模型, 并且求出最优解. (*提示*: 网络中的节点可以定义为 $[i, v]$, i 表示装入物品的序号, v 表示准备装入 i 之前背包剩余的容量.)

5. 一种老式的电子烤箱, 有底部相连的装有弹簧的门, 两扇门是方向相反向外开, 烤箱中间是加热元件. 通过用一只手打开烤箱门其中的一扇, 另一只手将面包片放入烤箱, 每次只能烘烤面包的一侧. 当面包片的一侧烘烤完毕后, 再将它翻转, 烘烤另外一边. 现在需要烘烤 3 片面包, 目标是确定这些操作 (放置、烘烤、翻转、取出) 的顺序使得烘烤这 3 片面包用的时间最少. 将这个问题转化成一个最短路径模型, 并利用下表中给出的各个操作时间的数据求解.

操　　作	所需时间 (秒)
将一片面包从任意一个门放入烤箱	3
烘烤面包的一侧	30
翻转烤箱中的面包	1
将一片面包从烤箱的任意一个门取出	3

6.3.2　最短路径算法

本节介绍求解有圈 (包含环) 网络和无圈网络的 2 个算法:

(1) Dijkstra 算法;　　(2) Floyd 算法.

Dijkstra 算法可以求网络中从源点到其他任何一个节点的最短路径. Floyd 算法的应用更加广泛一些, 它可以求网络中任意两个节点之间的最短路径.

Dijkstra 算法 用 u_i 表示从源点 1 到节点 i 的最短距离, 定义 $d_{ij}(\geqslant 0)$ 为弧 (i,j) 的长度, 那么算法根据这两个量立即可以给出节点 j 的标号为

$$[u_j, i] = [u_i + d_{ij}, i], \quad d_{ij} \geqslant 0$$

开始的节点标号为 $[0, -]$, 表示节点还没有父亲.

在 Dijkstra 算法中有两种类型的标号：暂时的和永久的. 对于暂时的标号而言, 如果这个节点找到了更短的路径, 那么标号就会改变. 如果找不到更短的路径, 那么这个标号就变成永久的.

第 0 步 对源点 (节点 1) 进行永久的标号 $[0, -]$. 令 $i = 1$.

第 1 步 (a) 计算节点 i 有边相连的每一个节点 j(并且 j 没有被永久标号) 的暂时标号 $[u_i + d_{ij}, i]$. 如果节点 j 已经通常另外一个节点 k 而被标号为 $[u_j, k]$, 并且 $u_i + d_{ij} < u_j$, 那么就用标号 $[u_i + d_{ij}, i]$ 代替 $[u_j, k]$.

(b) 如果所有的节点都得到了永久的标号, 那么停止. 否则, 从所有暂时标号的节点中选择具有最短距离 $(= u_r)$ 的标号 $[u_r, s]$. 令 $i = r$, 然后重复执行第 i 步.

例 6.3-4

图 6.15 中网络给出了从城市 1(节点 1) 和其他 4 个城市 (从节点 2 到节点 5) 之间可能的路线以及每条边的长度. 求城市 1 到其他 4 个城市的最短路径.

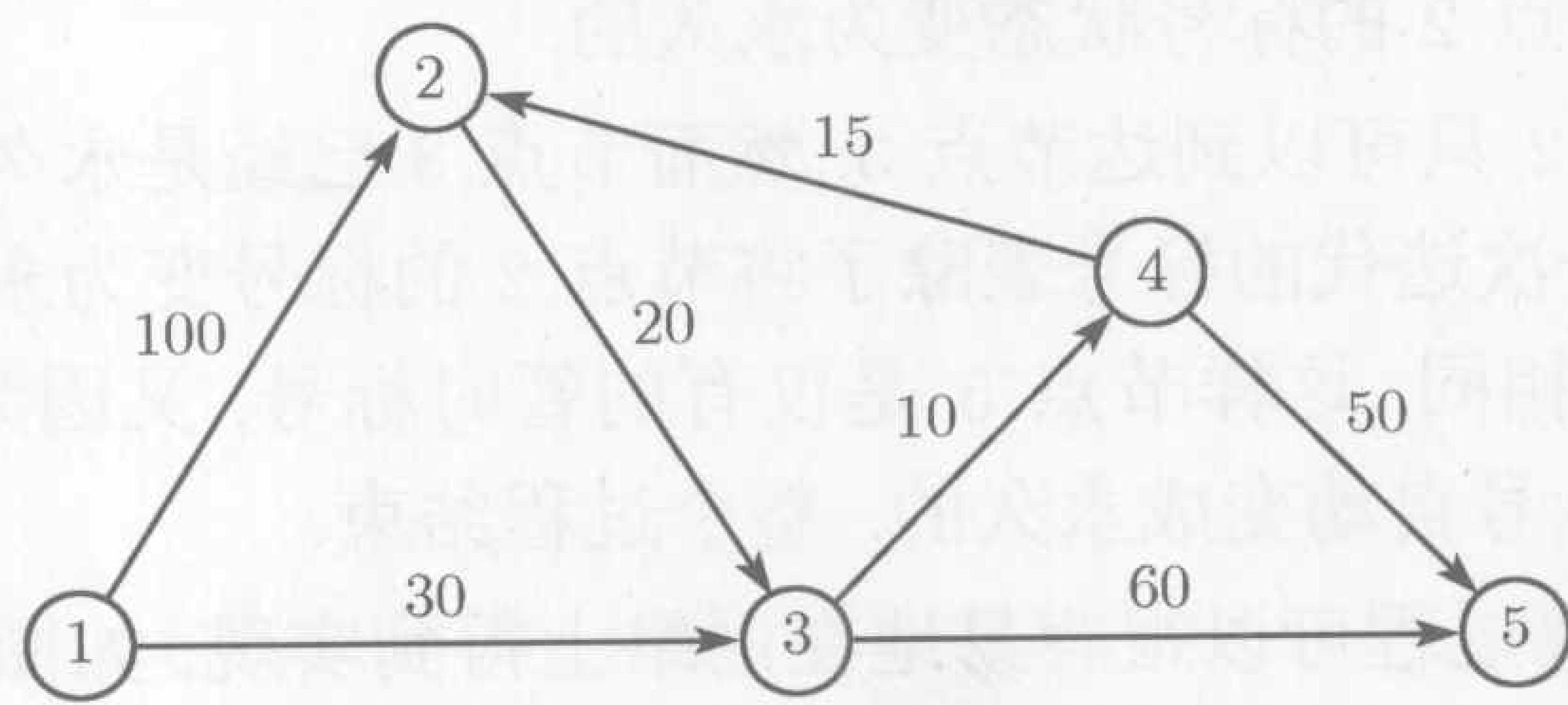


图 6.15 应用 Dijkstra 算法求解最短路径的网络例题

迭代 0 给节点 1 分配永久标号 $[0, -]$.

迭代 1 节点 1(永久标号的节点) 可以直接到达节点 2 和节点 3, 那么下表给出了节点新的标号 (暂时的和永久的):

节点	标号	标号的状态
1	**[0,—]**	**永久的**
2	$[0 + 100, 1] = [100, 1]$	暂时的
3	$[0 + 30, 1] = [30, 1]$	暂时的

表中有两个暂时的标号 $[100, 1]$ 和 $[30, 1]$, 节点 3 的距离较小 $(u_3 = 30)$. 因此

把节点 3 的标号状态转变为永久的.

迭代 2 节点 3 可以直接到达节点 4 和节点 5, 因此各个节点的标号变为:

节点	标号	标号的状态
1	[0,—]	永久的
2	[100, 1]	暂时的
3	**[30, 1]**	**永久的**
4	[30 + 10, 3] = [40, 3]	暂时的
5	[30 + 60, 3] = [90, 3]	暂时的

将暂时标号为 [40,3] 的节点 4 的标号状态转变为永久的 ($u_4 = 40$).

迭代 3 节点 4 可以直接到达节点 2 和节点 5, 那么各个节点的标号变为:

节点	标号	标号的状态
1	[0,—]	永久的
2	[40 + 15, 4] = [55, 4]	暂时的
3	[30, 1]	永久的
4	**[40, 3]**	**永久的**
5	[90, 3] 或者 [40 + 50, 4] = [90, 4]	暂时的

对于节点 2, 在迭代 1 得到的暂时标号为 [100, 1], 而此时通过节点 4 找到了一条更短的路径, 所以将节点 2 的标号更改为 [55, 4]. 同样在这次迭代后, 节点 5 有两个可以选择的标号, 并且标号对应的距离相等 ($u_5 = 90$). 通过比较, 这一次迭代选择节点 2, 并将节点 2 的标号状态变为永久的.

迭代 4 此时节点 2 只可以到达节点 3, 然而节点 3 已经是永久标号了, 不能给予重新标号. 所以这一次迭代的标号表除了将节点 2 的标号变为永久的以外, 其他的与迭代 3 的标号都相同. 这样节点 5 是仅有的暂时标号, 又因为节点 5 没有子节点, 所以节点 5 的标号自动变成永久的. 整个过程结束.

这个算法的计算过程可以很容易地在网络上得到实现, 如图 6.16 所示.

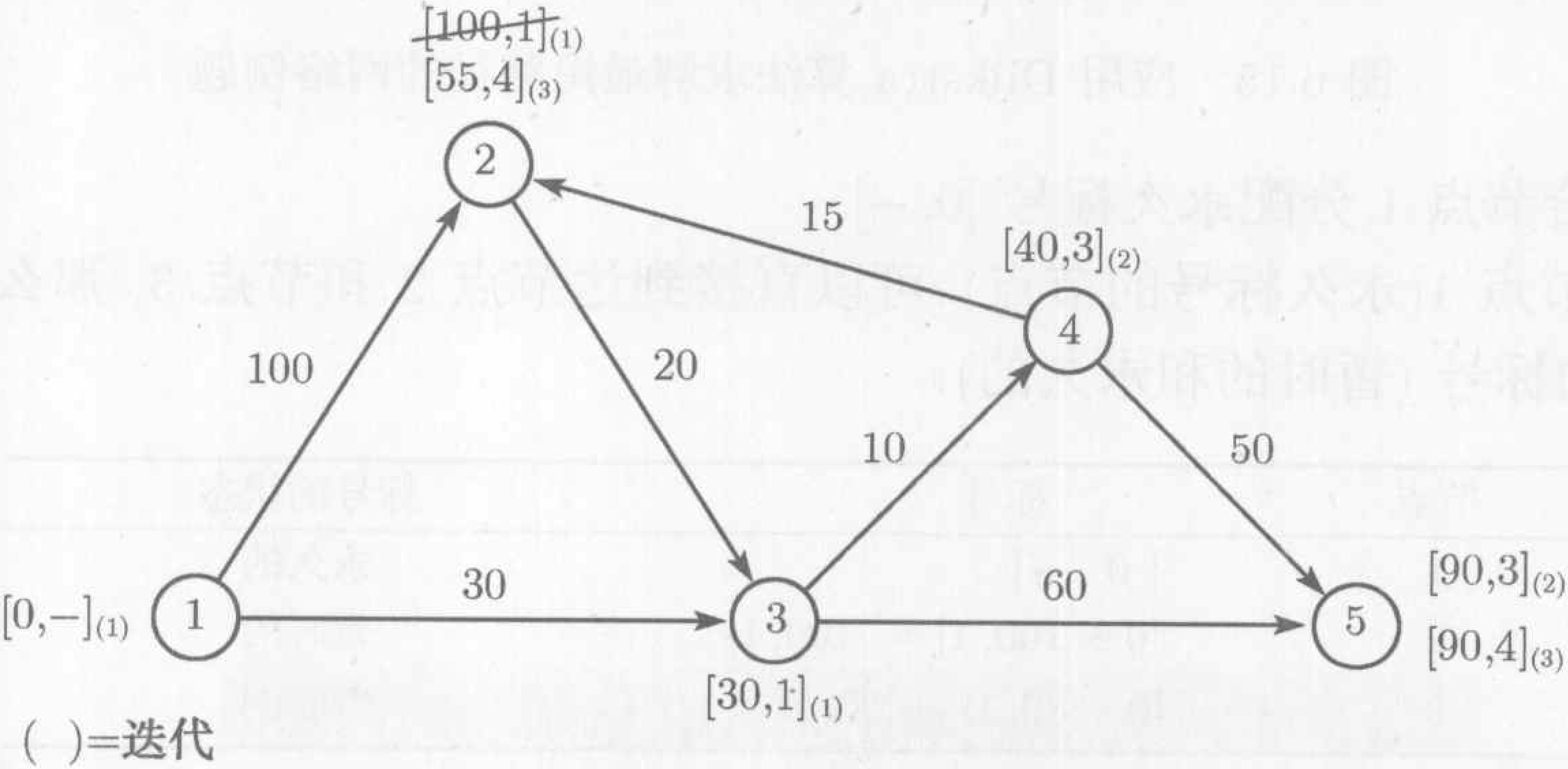


图 6.16 Dijkstra 算法的标号过程

从节点 1 到其他任何一个节点的最短路径可以通过这个节点的永久标号中的信息采用一步步回溯的方法得到. 例如, 节点 1 到节点 2 的最短路径可以通过下面的回溯方法找到:

$$(2) \to [55,4] \to (4) \to [40,3] \to (3) \to [30,1] \to (1)$$

那么所求的最短路径是 $1 \to 3 \to 4 \to 2$, 总长度为 55 英里.

TORA 程序

TORA 可以用于生成 Dijkstra 算法的迭代. 从 SOLVER/MODIFY 菜单中, 选择 Solve problem ⇒ Iterations ⇒ Dijkstra's algorithm. 文件 toraEx6.3-4. txt 中 TORA 使用的是例 6.3-4 的数据.

习题 6.3B

1. 图 6.17 中的网络给出了城市 $1, 2, \cdots, 8$ 之间的距离, 单位是英里. 应用 Dijkstra 算法, 求出下面指定的城市之间的最短路径.

(a) 城市 1 和城市 8; **(b)** 城市 1 和城市 6;

*__(c)__ 城市 4 和城市 8; **(d)** 城市 2 和城市 6.

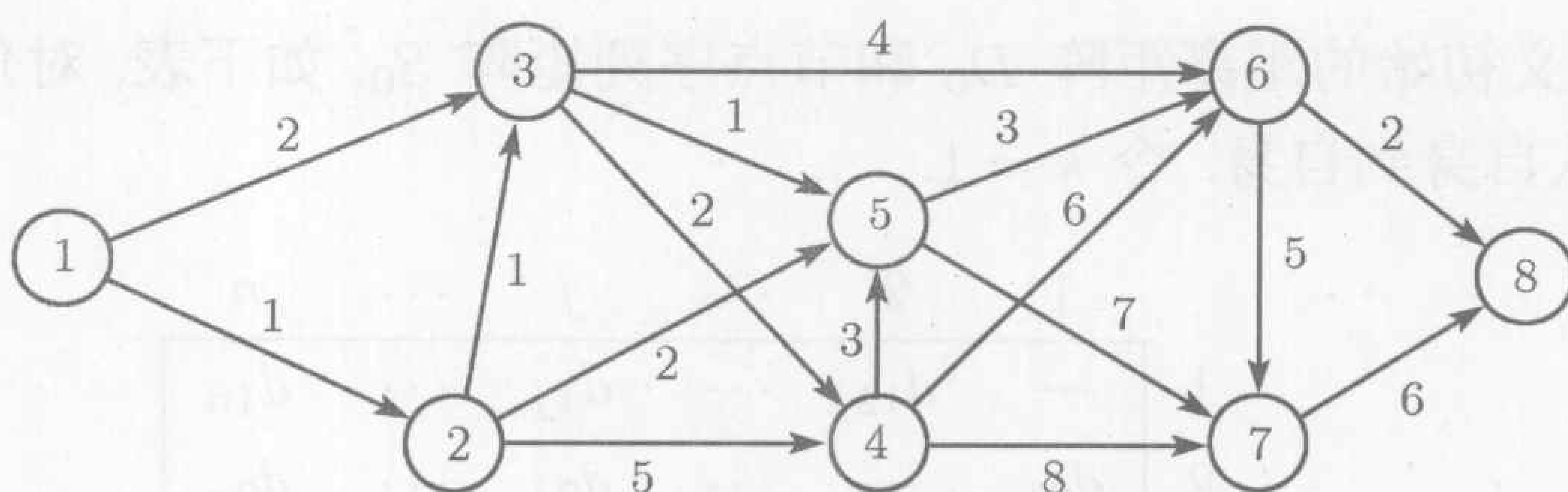


图 6.17 习题 6.3B 第 1 题中的网络

2. 应用 Dijkstra 算法, 求解图 6.18 网络中节点 1 到其他所有节点的最短路径.

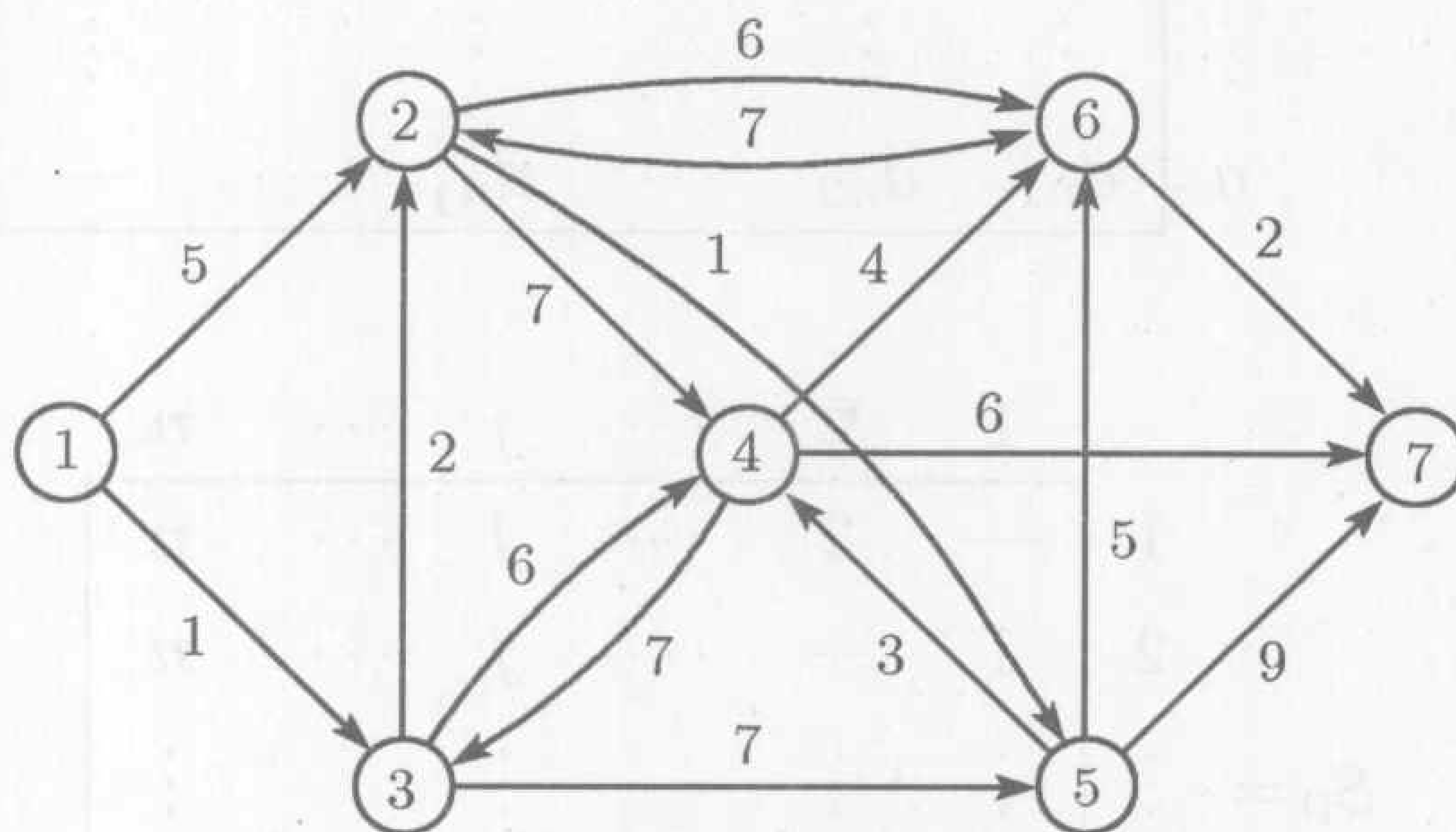


图 6.18 习题 6.3B 第 2 题中的网络

3. 应用 Dijkstra 算法求出下列问题的最优解:

(a) 习题 6.3A 中的第 1 题; **(b)** 习题 6.3A 中的第 2 题;

(c) 习题 6.3A 中的第 4 题.

Floyd 算法 Floyd 算法是比 Dijkstra 算法更一般的算法, 因为它可以给出网络中任意两个节点之间的最短路径. 算法首先将 n 个节点的网络表示成一个 n 行 n 列的矩阵, 矩阵中的元素 (i,j) 表示从节点 i 到节点 j 的距离 d_{ij}, 如果 i,j 之间没有边相连, 那么相应的元素就是无穷.

Floyd 算法的思想很直观. 如图 6.19 所示, 给定 3 个节点 i, j, k, 以及它们之间的距离, 如果满足

$$d_{ik}+d_{kj}<d_{ij}$$

那么从 i 经过 k 到 j 更短. 在这种情况下, 用间接的路径 $i\to k\to j$ 代替路径 $i\to j$. 下面的步骤给出了这种**三重操作**替换应用在网络上的具体算法.

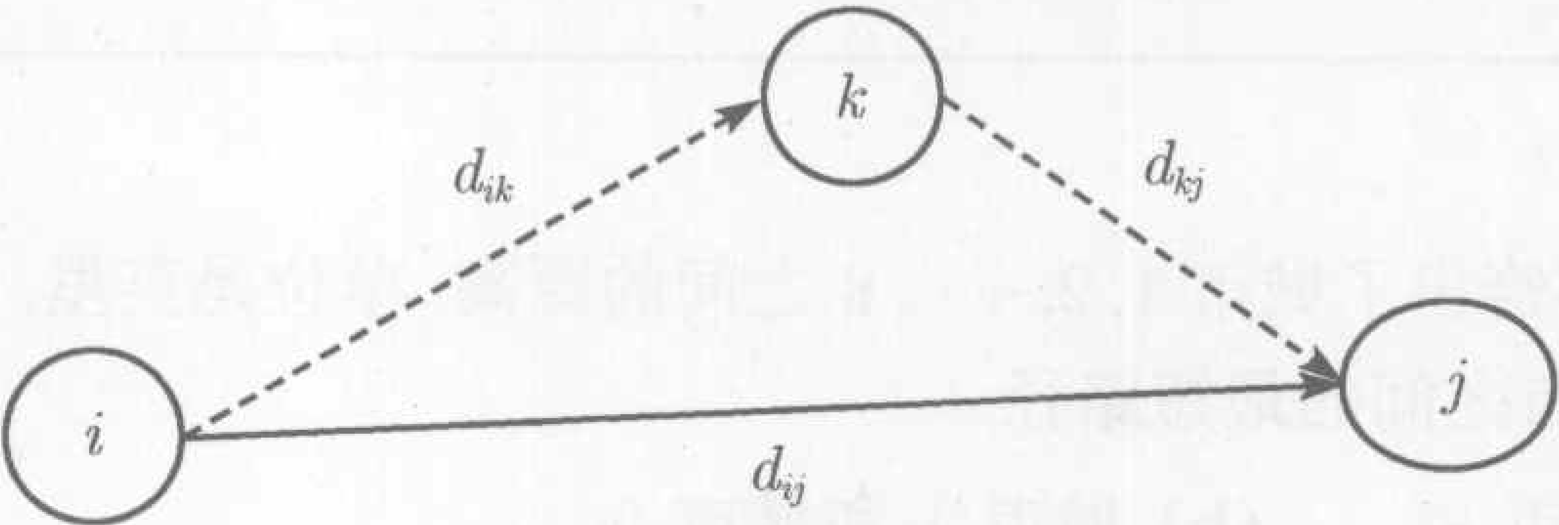


图 6.19 Floyd 三重操作

第 0 步 定义初始的距离矩阵 D_0 和节点序列矩阵 S_0, 如下表, 对角线上用 (—) 表示不必要从自身到自身. 令 $k=1$.

$$D_0=\begin{array}{c|cccccc|}
 & 1 & 2 & \cdots & j & \cdots & n \\
\hline
1 & — & d_{12} & \cdots & d_{1j} & \cdots & d_{1n} \\
2 & d_{21} & — & \cdots & d_{2j} & \cdots & d_{2n} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
i & d_{i1} & d_{i2} & \cdots & d_{ij} & & d_{in} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
n & d_{n1} & d_{n2} & \cdots & d_{nj} & \cdots & — \\
\hline
\end{array}$$

$$S_0=\begin{array}{c|cccccc|}
 & 1 & 2 & \cdots & j & \cdots & n \\
\hline
1 & — & 2 & \cdots & j & \cdots & n \\
2 & 1 & — & \cdots & j & \cdots & n \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
i & 1 & 2 & & j & & n \\
\vdots & \vdots & \vdots & & & & \vdots \\
n & 1 & 2 & \cdots & j & \cdots & — \\
\hline
\end{array}$$

一般的第 k 步 令第 k 行和第 k 列为枢轴行和枢轴列. 对于矩阵 D_{k-1} 中的每一个元素 d_{ij} 做三重操作. 如果满足条件:

$$d_{ik} + d_{kj} < d_{ij} \quad (i \neq k, j \neq k, i \neq j)$$

那么进行下面的转化:

(a) 用 $d_{ik} + d_{kj}$ 代替矩阵 D_{k-1} 中的元素 d_{ij}, 从而得到矩阵 D_k.

(b) 用 k 代替矩阵 S_{k-1} 中的元素 s_{ij}, 从而得到矩阵 S_k. 令 $k = k+1$. 如果 $k = n+1$, 停止; 否则, 重复第 k 步.

图 6.20 形象地给出了算法第 k 步对矩阵 D_{k-1} 的转化过程, 其中第 k 行和第 k 列是当前的枢轴行和枢轴列. 第 i 行表示第 $1, 2, \cdots, k-1$ 行中的任何一个, 第 p 行表示第 $k+1, k+2, \cdots, n$ 行中的任何一个, 第 j 列表示第 $1, 2, \cdots, k-1$ 行中的任何一个, 第 q 列表示第 $k+1, k+2, \cdots, n$ 行中的任何一个. 三重操作可以按照下面的方法执行: 如果枢轴行和枢轴列上的元素 (图中正方形中的元素) 之和小于相应交叉元素 (图中圆形中的元素) 的值, 那么用这个和代替交叉元素的值就可以使最短距离得到优化.

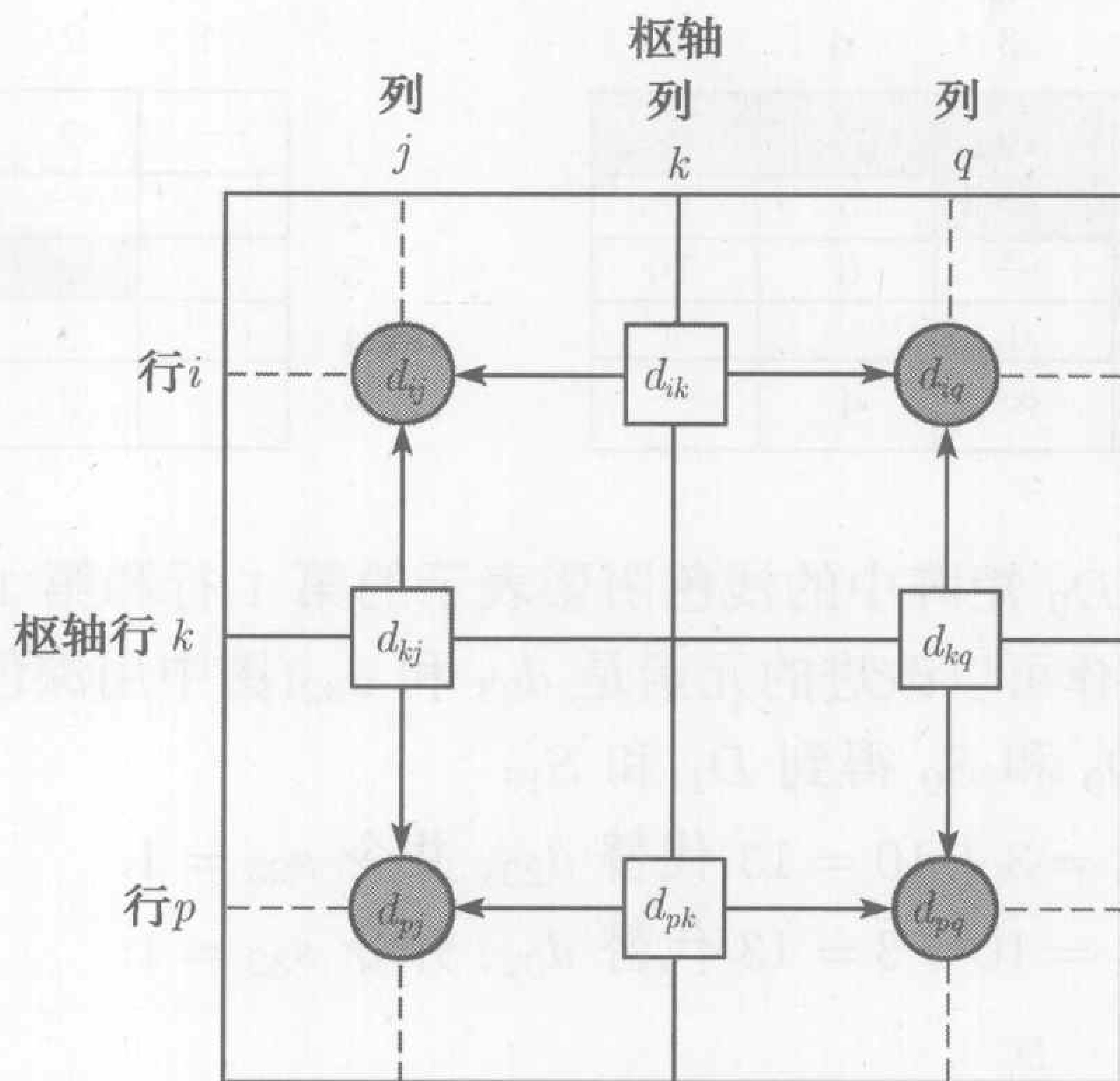


图 6.20 在矩阵中实现三重操作

经过 n 步之后, 我们可以从矩阵 D_n 和 S_n 中按照下面的规则得到节点 i 和 j 之间的最短路径:

(1) 在矩阵 D_n 中, d_{ij} 表示节点 i 与节点 j 之间的最短路径长度.

(2) 在矩阵 S_n 中, 可以确定中间节点 $k = s_{ij}$(根据得到的路径 $i \to k \to j$). 如果 $s_{ik} = k$ 并且 $s_{kj} = j$, 停止, 因为路径中所有的中间节点都找到了. 否则, 在节点 i 与节点 k 之间, 节点 k 与节点 j 之间重复上面的程序.

例 6.3-5

对于图 6.21 中给出的网络, 求任意两个节点之间的最短路径. 图中弧上给出了相应节点间的距离 (单位: 英里). 弧 (3,5) 是有向的, 因此不允许从节点 5 到节点 3. 其他边都是双向的.

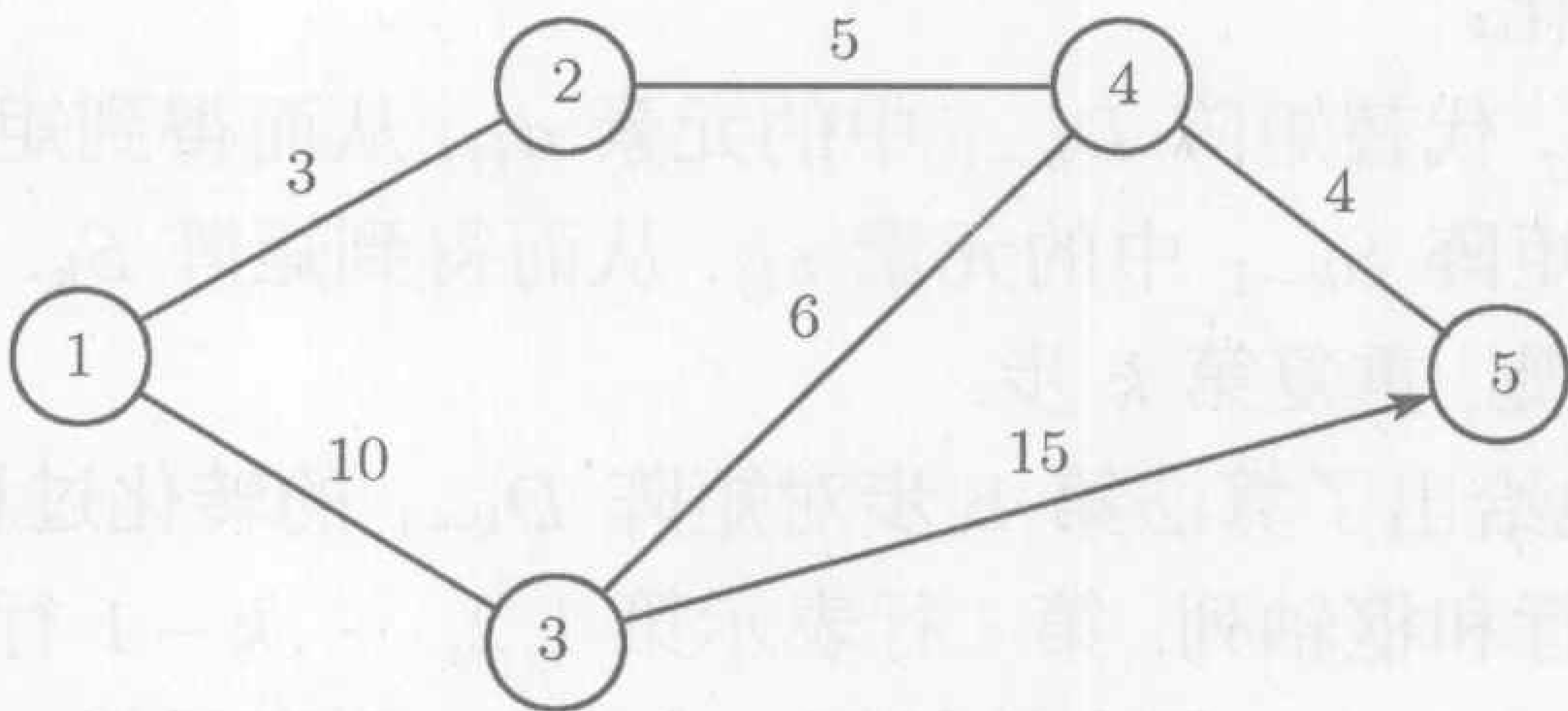


图 6.21 例 6.3-5 中的网络

迭代 0 矩阵 D_0 和 S_0 代表初始的网络. 除了 $d_{53}=\infty$ 外 (因为不允许从节点 5 到节点 3), D_0 是对称的.

D_0

	1	2	3	4	5
1	—	3	10	∞	∞
2	3	—	∞	5	∞
3	10	∞	—	6	15
4	∞	5	6	—	4
5	∞	∞	∞	4	—

S_0

	1	2	3	4	5
1	—	2	3	4	5
2	1	—	3	4	5
3	1	2	—	4	5
4	1	2	3	—	5
5	1	2	3	4	—

迭代 1 令 $k=1$. D_0 矩阵中的浅色阴影表示的第 1 行和第 1 列为枢轴行和枢轴列, 其中应用三重操作可以改进的元素是 d_{23} 和 d_{32}(图中用深色阴影表示). 然后根据下面的方法, 从 D_0 和 S_0 得到 D_1 和 S_1:

(1) 用 $d_{21}+d_{13}=3+10=13$ 代替 d_{23}, 并令 $s_{23}=1$.

(2) 用 $d_{31}+d_{12}=10+3=13$ 代替 d_{32}, 并令 $s_{32}=1$.

D_1

	1	2	3	4	5
1	—	3	10	∞	∞
2	3	—	**13**	5	∞
3	10	**13**	—	6	15
4	∞	5	6	—	4
5	∞	∞	∞	4	—

S_1

	1	2	3	4	5
1	—	2	3	4	5
2	1	—	**1**	4	5
3	1	**1**	—	4	5
4	1	2	3	—	5
5	1	2	3	4	—

迭代 2 令 $k=2$. D_1 矩阵中的浅色阴影表示的第 2 行和第 2 列为枢轴行和枢轴列, 其中应用三重操作可以改进的元素是 d_{14} 和 d_{41}(图中用深色阴影表示). 结果变

化为 D_2 和 S_2 中的粗体元素:

D_2

	1	2	3	4	5
1	—	3	10	8	∞
2	3	—	13	5	∞
3	10	**13**	—	6	15
4	8	5	6	—	4
5	∞	∞	∞	4	—

S_2

	1	2	3	4	5
1	—	2	3	**2**	5
2	1	—	1	4	5
3	1	1	—	4	5
4	**2**	2	3	—	5
5	1	2	3	4	—

迭代 3 令 $k = 3$, D_2 矩阵中的浅色阴影表示的第 3 行和第 3 列为枢轴行和枢轴列. 应用三重操作可以得到矩阵 D_3 和 S_3:

D_3

	1	2	3	4	5
1	—	3	10	8	**25**
2	3	—	13	5	**28**
3	10	13	—	6	15
4	8	5	6	—	4
5	∞	∞	∞	4	—

S_3

	1	2	3	4	5
1	—	2	3	2	**3**
2	1	—	1	4	**3**
3	1	1	—	4	5
4	2	2	3	—	5
5	1	2	3	4	—

迭代 4 令 $k = 4$, D_3 矩阵中的浅色阴影表示的第 4 行和第 4 列为枢轴行和枢轴列. 应用三重操作可以得到矩阵 D_4 和 S_4:

D_4

	1	2	3	4	5
1	—	3	10	8	**12**
2	3	—	**11**	5	**9**
3	10	**11**	—	6	**10**
4	8	5	6	—	4
5	**12**	**9**	**10**	4	—

S_4

	1	2	3	4	5
1	—	2	3	2	**4**
2	1	—	**4**	4	**4**
3	1	**4**	—	4	**4**
4	2	2	3	—	5
5	**4**	**4**	**4**	4	—

迭代 5 令 $k = 5$, D_4 矩阵中的浅色阴影表示的第 5 行和第 5 列为枢轴行和枢轴列. 应用三重操作发现没有可以改进的元素了.

最后得到的矩阵 D_4 和 S_4 包含了网络中任意两个节点最短路径的所有信息. 例如, 从矩阵 D_4 中可以得到节点 1 与节点 5 之间的最短路径长度是 $d_{15} = 12$ 英里. 下面通过 S_4 确定这条最短路径, 注意到对于一个路段 (i, j), 如果 $s_{ij} = j$, 那么从 i 有弧直接连接到 j; 否则, i 和 j 之间经过至少一个中间节点. 在 S_4 中, 因为 $s_{15} = 4 \neq 5$, 可以知道路径应该经过节点 4 到达节点 5, 即 $1 \to 4 \to 5$; 又因为 $d_{14} = 2 \neq 4$ 所以 $(1, 4)$ 也不是直接相连, 应该用路径 $1 \to 2 \to 4$ 代替, 因此路径 $1 \to 4 \to 5$ 变为 $1 \to 2 \to 4 \to 5$. 此时 $s_{12} = 2, s_{24} = 4, s_{45} = 5$, 所以这条最短路径

就是 $1 \to 2 \to 4 \to 5$.

TORA 程序

与应用在 Dijkstra 算法中类似, TORA 也可以用来执行 Floyd 算法的迭代. 从 SOLVER/MODIFY 菜单, 选择 Solve problem ⇒ Iterations ⇒ Floyd's algorithm. 文件 toraEx6.3-5. txt 中 TORA 使用的是例 6.3-5 的数据.

习题 6.3C

1. 在例 6.3-5 中, 应用 Floyd 算法确定下面每一对节点之间的最短路径:

*(a) 从节点 5 到节点 1; (b) 从节点 3 到节点 5;

(c) 从节点 5 到节点 3; (d) 从节点 5 到节点 2.

2. 对于图 6.22 给定的网络, 应用 Floyd 算法确定下面每一对节点之间的最短路径. 其中弧 $(7,6)$ 和 $(6,4)$ 是单向的, 图中表示的距离单位是英里.

(a) 从节点 1 到节点 7; (b) 从节点 7 到节点 1;

(c) 从节点 6 到节点 7.

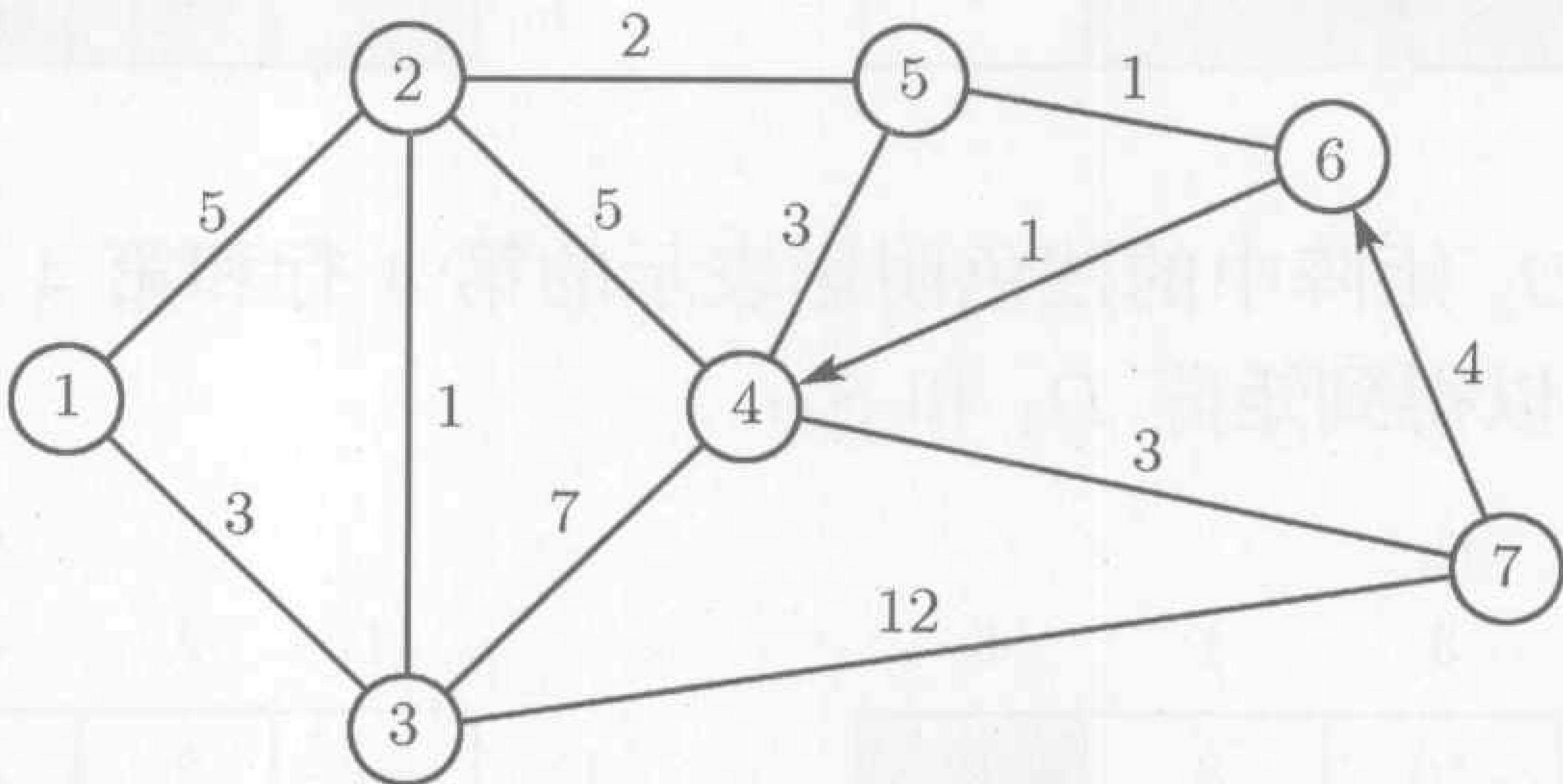


图 6.22 习题 6.3C 第 2 题对应的网络

3. Tell-All 移动电话公司需要为 6 个地理区域提供服务, 图 6.23 给出了 6 个区域之间的卫星距离 (英里). Tell-All 公司希望确定任意两个区域之间的最短路径, 使得信息传输更加快速有效.

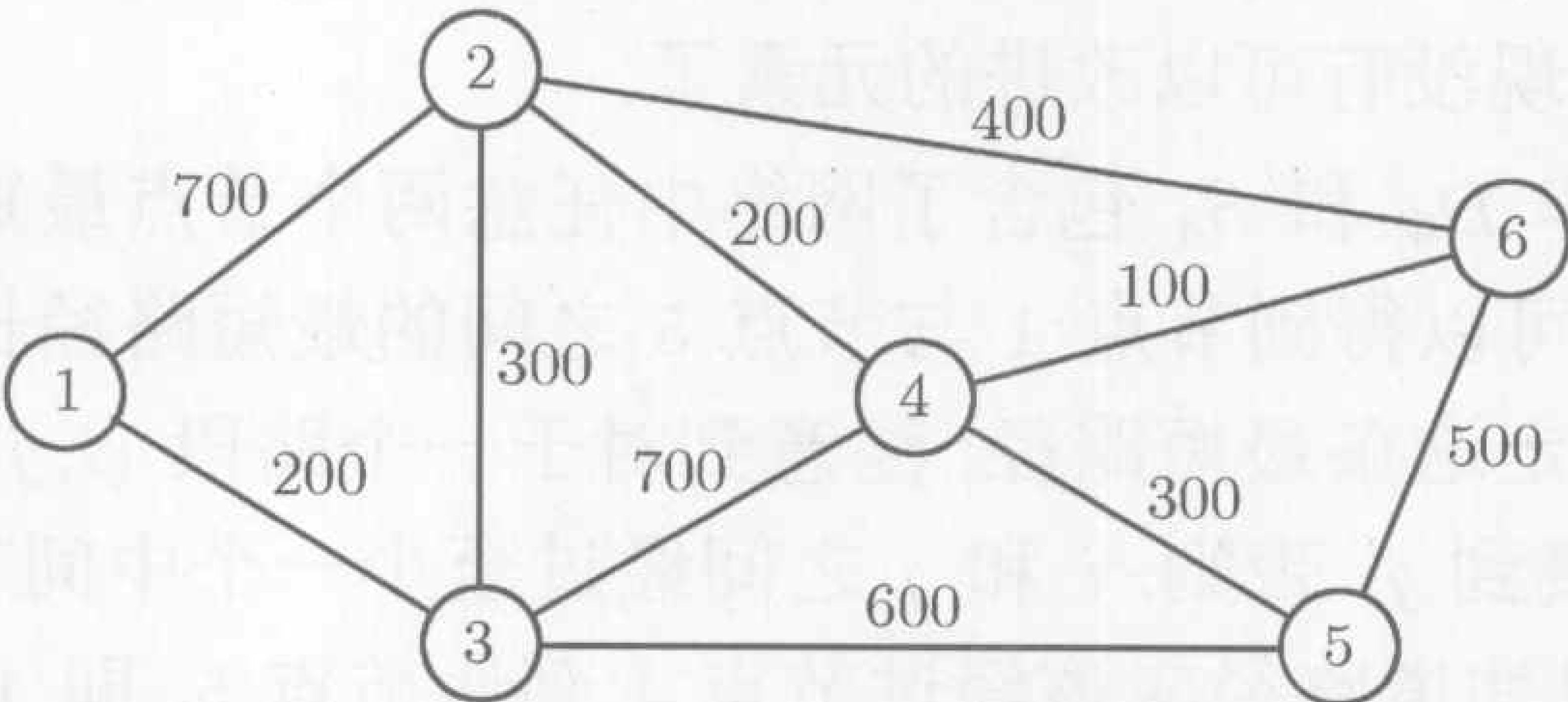


图 6.23 习题 6.3C 第 3 题对应的网络

***4.** 6 个小朋友 Joe, Kay, Jim, Bob, Rae 和 Kim 一起玩一个新的捉迷藏游戏, 一个小朋友的藏身之处只有其他一部分小朋友知道. 当一个小朋友需要找到另一个小孩的藏身之处时, 如果他不知道藏在哪儿, 那么需要先找到另外的小朋友, 然后得到相关的信息, 因此, 在找到这个小朋友时, 他或许需要经过多个其他的小朋友. 例如, 假定 Joe 需要找到 Kim, Joe 知道 Jim 藏在哪儿, 而 Jim 知道 Kim 藏在哪儿, 那么, Joe 需要先找到 Jim, 然后 Jim 告诉他 Kim 的藏身之处. 下面给出了相关的信息:

Joe 知道 Bob 和 Kim 的藏身之处;
Kay 知道 Bob, Jim 和 Rae 的藏身之处;
Jim 和 Bob 都只知道 Kay 藏身之处;
Rae 知道 Kim 的藏身之处;
Kim 知道 Joe 和 Bob 的藏身之处.

给每一个小朋友设计一个寻找其他任意一个小朋友最快的方案, 即经过最少的接触就可以找到所有其他的小朋友. 对于接触次数最多的方案又是怎样?

6.3.3 最短路径问题的线性规划模型

本节介绍最短路径问题的线性规划模型, 应用这个模型可以求出网络中任意两个节点之间的最短路径, 从这点看, 它与 Floyd 算法一致.

假定求最短路径的网络包含 n 个节点, 需要求出网络中任意两个节点 s 和 t 之间的最短路径. 线性规划模型假定有一个单位的流从节点 s 流入网络, 从节点 t 流出.

定义

$$\begin{aligned} x_{ij} &= \text{弧}(i,j)\text{上的流量} \\ &= \begin{cases} 1, & \text{如果弧}(i,j)\text{在最短路径上} \\ 0, & \text{如果弧}(i,j)\text{不在最短路径上} \end{cases} \\ c_{ij} &= \text{弧}(i,j)\text{的长度} \end{aligned}$$

那么, 线性规划的目标函数是

$$\min z = \sum_{\text{弧}(i,j)} c_{ij} x_{ij}$$

每一个节点上的约束表示为*流守恒方程*:

$$\text{总的流入量} = \text{总的流出量}$$

对于节点 j, 用数学的语言表示上面的等式就是

$$\begin{pmatrix} \text{外部输入到} \\ \text{节点 } j \text{ 的流量} \end{pmatrix} + \sum_{\substack{i \\ \text{弧}(i,j)}} x_{ij} = \begin{pmatrix} \text{从节点 } j \text{ 外部} \\ \text{输出的流量} \end{pmatrix} + \sum_{\substack{k \\ \text{弧}(j,k)}} x_{jk}$$

例 6.3-6

考虑例 6.3-4 中求最短路径的网络. 现在我们需要确定节点 1 到节点 2 的最短路径, 也就是 $s=1, t=2$. 图 6.24 表示出一个单位的流是如何从节点 1 流入, 以及从节点 2 流出.

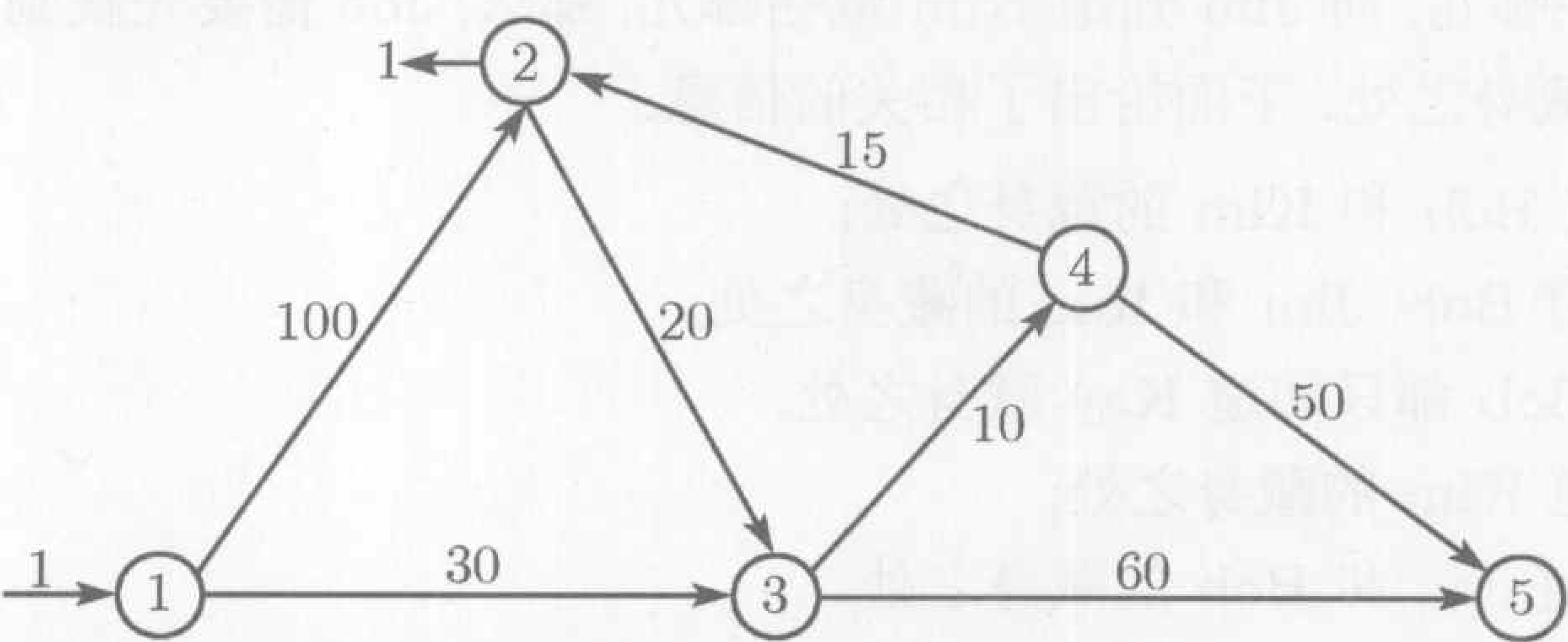


图 6.24　通过输入一个单位的流来确定节点 $s=1$ 与节点 $t=2$ 之间的最短路径

从网络中可以得出每个节点的流守恒方程:

$$\begin{aligned}
&\text{节点 1:} \quad 1 = x_{12} + x_{13} \\
&\text{节点 2:} \quad x_{12} + x_{42} = x_{23} + 1 \\
&\text{节点 3:} \quad x_{13} + x_{23} = x_{34} + x_{35} \\
&\text{节点 4:} \quad x_{34} = x_{42} + x_{45} \\
&\text{节点 5:} \quad x_{35} + x_{45} = 0
\end{aligned}$$

完整的线性规划模型可以表示为

	x_{12}	x_{13}	x_{23}	x_{34}	x_{35}	x_{42}	x_{45}		
min $z=$	100	30	20	10	60	15	50		
节点 1	1	1						=	1
节点 2	−1		1			−1		=	−1
节点 3		−1	−1	1	1			=	0
节点 4				−1		1	1	=	0
节点 5					−1		−1	=	0

根据网络线性规划模型的一个典型性质, x_{ij} 列中恰好有一个“1”和一个“−1”, 分别位于第 i 行和第 j 行.

最优解是 (应用 TORA 求解, 见文件 toraEx6.3-6.txt)

$$z = 55, x_{13} = 1, x_{34} = 1, x_{42} = 1$$

这个解对应的从节点 1 到节点 2 的最短路径是 $1 \to 3 \to 4 \to 2$, 相应的长度是 $z = 55$(英里).

习题 6.3D

1. 在例 6.3-6 中, 应用线性规划模型确定下面节点对之间的最短路径长度:

*(a) 节点 1 到节点 5;　　(b) 节点 2 到节点 5.

Excel 规划求解程序

图 6.25 给出了应用 Excel 规划求解电子表格解例 6.3-6 中从始点 N1 到终点 N2 的最短路径 (见文件 solverEx6.3-6.xls). 模型的输入数据是 B3:B6 单元格中的距离矩阵. 注意到矩阵中没有节点 N1 对应的列, 这是因为节点 N1 没有入弧, 同样节点 N5 没有行是因为它没有出弧. 矩阵中的空白单元格表示图中没有相应的弧. 通过在单元格 F3 和 B7 中输入 1 表示节点 N1 和 N2 分别是始点和终点. 如需要求解其他节点对之间的最短路径, 只需要将相应的单元格中输入 1. 例如, 求解节点 N2 到节点 N4 的最短路径, 那么在单元格 F4 和 D7 中输入 1 即可.

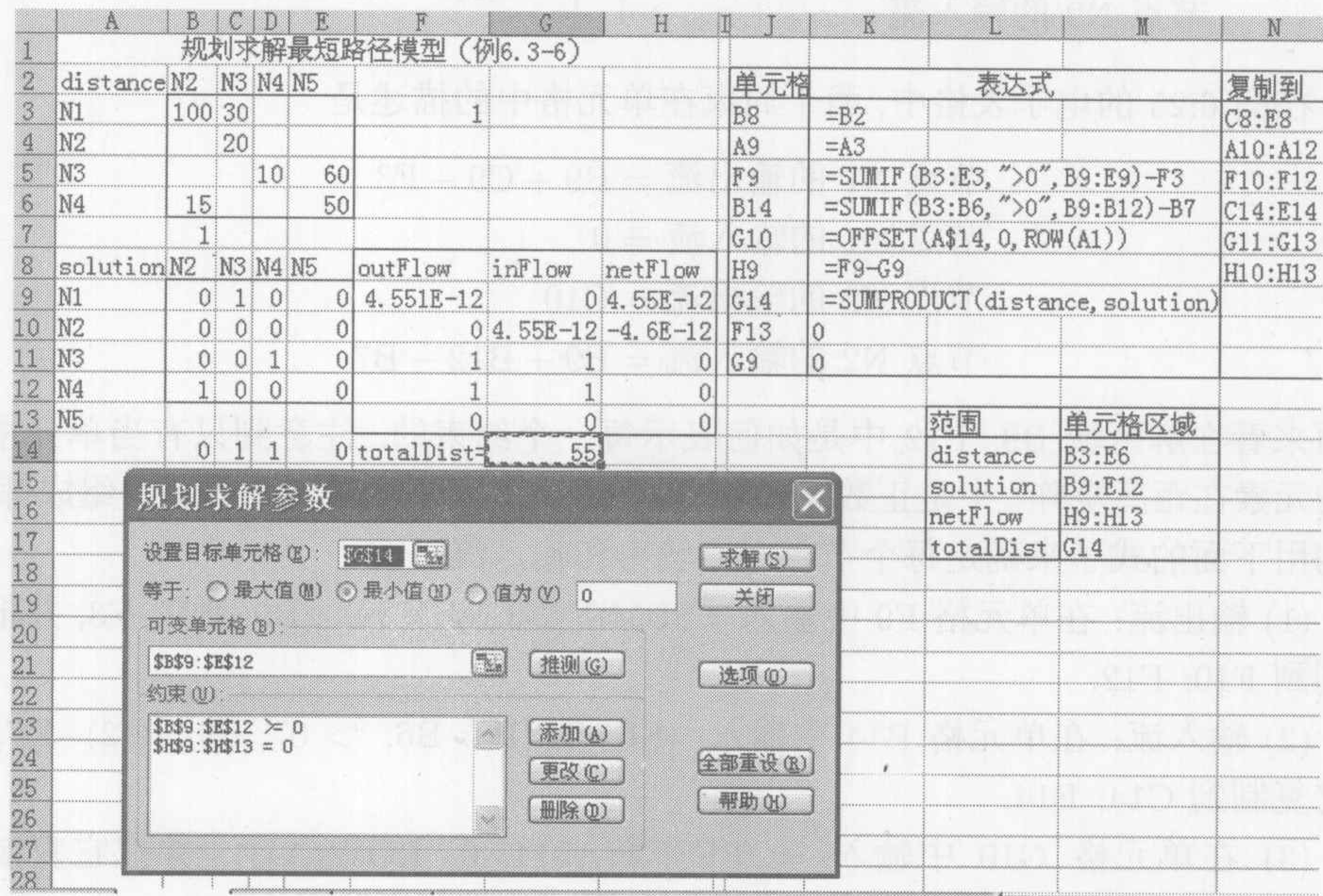


图 6.25 Excel 规划求解计算例 6.3-6 中节点 1 与节点 2 之间的最短路径 (见文件 solverEx6.3-6.xls)

程序输出的解在 B9:E12 的单元格中. 每一个单元格对应了连接两个节点的弧, 例如单元格 C10 对应了弧 (N2, N3), 相应的变量是 x_{23}. 如果一条弧 (Ni, Nj) 位于最短路径上, 那么单元格变量 $x_{ij}=1$, 否则为 0.

根据 B3:E6 的距离矩阵 (命名为distance) 和 B9:E12 的解矩阵 (命名为solution), 目标函数的计算结果在单元格 G14 中输入 =SUMPRODUCT(B3:E6, B9:E12), 或者 =SUMPRODUCT(distance, solution). 读者或许很想知道在距离矩阵中为什么使用了空白输入 (Excel 表格中没有输入 0) 以及这对目标函数会有什么影响. 这一点我们将会在接下来加以解释, 首先来看问题约束中的变量是如何消除的.

正如在例 6.3-6 中线性规划模型中解释的那样, 问题的约束有一个一般的形式:

$$\text{流出量} - \text{流入量} = 0$$

根据这个定义可以容易地将节点的外部单位流 (如果存在的话) 合并到流出量或者流入量方程中. 例如, 在例 6.3-6 中, 外部单位流从节点 N1 流入, 从 N2 流出, 那么相应的约束是

$$\left.\begin{array}{l}\text{节点 N1 的输出流} = x_{12} + x_{13} - 1\\ \text{节点 N1 的输入流} = 0\end{array}\right\} \Rightarrow x_{12} + x_{13} - 1 = 0$$

$$\left.\begin{array}{l}\text{节点 N2 的输出流} = x_{23}\\ \text{节点 N2 的输入流} = x_{12} + x_{42} - 1\end{array}\right\} \Rightarrow x_{23} - x_{12} - x_{42} - 1 = 0$$

在图 6.25 的电子表格中, 两个约束在单元格中的描述是

节点 N1 的输出流 = B9 + C9 − F3
节点 N1 的输入流 = 0
节点 N2 的输出流 = C10
节点 N2 的输入流 = B9 + B12 − B7

下面来看在解区域 B9: E12 中是如何表示每一个约束的. 注意到只有当单元格包含的元素在距离矩阵① 中是正数时, 这个单元格中的元素才是一个约束的组成部分. 我们用下面的式子来确定每个节点上的输出和输入流:

(1) 输出流: 在单元格 F9 中输入 =SUMIF(B3:E3, “ >0”, B9: E9)−F3, 并把它复制到 F10: F12.

(2) 输入流: 在单元格 B14 中输入 =SUMIF(B3: B6, “> 0”, B9: B12)−B7, 并把它复制到 C14: E14.

(3) 在单元格 G10 中输入 =OFFSET(A$14, 0, ROW(A1)), 并把它复制到 G11: G13 来调换 G 列的输入流.

(4) 在单元格 G9 和 F13 中输入 0, 表示节点 N1 没有输入流以及节点 N5 没有输出流 (依照电子表格的定义).

(5) 在单元格 H9 中输入 =F9−G9, 并把它复制到 H10: H13 中, 然后计算网络流.

现在这张电子表格已经可以应用 Excel 规划求解, 如图 6.25 所示. 有一件非常奇怪的事情: 当在**规划求解参数**对话框中输入outFlow = inFlow来定义约束时, 规划求解不能给出可行解, 甚至是调整**规划求解选项**对话框中的精度也不行. (经验告诉我们, 此时解区域 B9:E12 必须重新设置 0 或者空白.) 更奇怪的是, 如果约束用inFlow=outFlow代替, 最优解就可以找到. 在文件 solverEx6.3-6.xls 中, 使

① 如果一个问题碰巧出现两个节点的距离是 0, 那么用一个非常小的正数代替它.

用netFlow区域 H9: H12, 用netFlow = 0 表示没有问题的约束. 虽然还不是很清楚为什么会这样, 但问题可能与舍入误差有关.

图 6.25 的输出结果是 (N1−N3=1, N3−N4=1, N4−N2=1), 总的距离是 55 英里, 可以得到最优路径是 $1 \to 3 \to 4 \to 2$.

评注 在许多参考书中, 网络都是通过明确的弧 (节点 i、节点 j, 以及距离) 定义的, 当网络中弧的数目很多的时候, 计算量就会很大. 然而我们的模型是用距离矩阵表示网络, 并且容易表示所有需要的流约束. 有人或许会提出疑问, 在模型中使用了 $(n-1)\times(n-1)$ 个 x_{ij} 变量, 这样表示需要的变量个数要比使用弧表示网络的模型的变量数目多 (例如, 例 6.3-6 中有 7 条弧, 也就是 7 个变量, 然而我们的模型需要 $4\times 4=16$ 个变量). 要注意的是这些额外的变量只会出现在目标函数中, 而且相应的系数是 0(输入是空白), 并且流约束的表示也和其他的恰好一样(每一个SUMIF函数). 因此, 在商用软件中的预处理时都可以发现这些"现象"并且在使用单纯形法之前就可以自动地去除这些额外变量, 所以不会带来额外的计算量.

AMPL 程序

图 6.26 给出了求解例 6.3-6 的 AMPL 程序 (文件 amplEx6.3-6a.txt). 如果弧[i,j]在最短路径上, 那么变量x[i,j]取值为 1, 否则取值为 0. 不管问题的规模多少, 一般地这个模型可以用于求解网络中任意两个节点之间的最短路径.

正如例 6.3-6 中说明的, AMPL 将问题看作是在一个网络中从一个特殊的节点start流入网络一个单位的流, 从节点end流出. 模型主要的数据输入就是一个表示连接节点 i 和节点 j 的弧的长度d[i,j]的 $n\times n$ 的矩阵. 按照 AMPL 的语法, 在矩阵中, 用点作为占位府表示这个位置上对应的弧没有长度. 在模型中, 使用语句param d{i in 1..n, j in 1..n} default M; 表示给矩阵中元素代表的距离赋值为无穷大的数M(=999 999), 这就使得相应的弧的长度为无穷大. 也可以直接在矩阵中将相应的元素用 999 999 代替得到同样的效果, 但是这样会导致数据输入比较混乱, 很不方便.

在每个节点上表示流守恒的约束是:

$$流入量 - 流出量 = 外部流量$$

从x[i,j], 我们可以利用下面的语句定义节点 i 上的输入流和输出流:

```
var inFlow{j in 1..n} = sum{i in 1..n}x[i,j];
var outFlow{i in 1..n} = sum{j in 1..n}x[i,j];
```

第 i 个约束的左端项就是outFlow[i]-inFlow[i].

第 i 个约束的右边 (节点 i 上的外部流) 定义为

```
param rhs{i in 1..n}=if i=start then 1 else(if i=end then -1 else 0);
```

(参看附录 A.3 了解语句if then else) 根据这个语句, 指定节点start和end自动的赋值 1, −1 或者 0 到约束右侧的rhs.

```
#---- shortest route model (Example 6.3-6) ----
param n;
param start;
param end;
param M=999999; #infinity
param d{i in 1..n, j in 1..n} default M;
param rhs{i in 1..n}=if i=start then 1
                   else (if i=end then -1 else 0);
var x{i in 1..n,j in 1..n}>=0;
var outFlow{i in 1..n}=sum{j in 1..n}x[i,j];
var inFlow{j in 1..n}=sum{i in 1..n}x[i,j];

minimize z: sum{i in 1..n, j in 1..n}d[i,j]*x[i,j];
subject to limit{i in 1..n}:outFlow[i]-inFlow[i]=rhs[i];

data;
param n:=5;
param start:=1;
param end:=2;
param d:
   1 2   3  4  5:=
1 . 100 30 .  .
2 . .   20 .  .
3 . .   .  10 60
4 . 15  .  .  50
5 . .   .  .  .;
solve;
print "Shortest length from",start,"to",end,"=",z;
printf "Associated route: %2i",start;
for {i in 1..n-1} for {j in 2..n}
     {if x[i,j]=1 then printf "- %2i", j;} print;
```

图 6.26　AMPL 最短路径模型

目标函数是对所有的d[i,j]*x[i,j]之和求最小值.

在当前的例题中, start=1和end=2, 表示需要求节点 1 到节点 2 的最短路径. 相应的输出是

```
Shortest length from 1 to 2 = 55
Associated route: 1 - 3 - 4 - 2
```

评注　图 6.26 给出的 AMPL 模型有一个缺点: 活动变量 x_{ij} 的个数是 n^2 个, 这显然比网络中真实的弧 (具有正的距离的弧) 的数目大得多, 导致问题的规模很大. 这是因为在模型中我们将不存在的弧的距离赋值为一个无穷大的数M(= 999 999), 用来确保在最优解中这些相应的活动变量取值为零. 可以使用下面的语句来弥补这个缺点:

```
var x{i in 1..n,j in 1..n:d[i,j]<M}>=0;
```

(参看附录 A.4 了解这个语句的用法) 通过使用下面的语句来对约束应用同样的方法:

```
var inFlow{j in 1..n}=sum{i in 1..n:d[i,j]<M}x[i,j];
var outFlow{i in 1..n}=sum{j in 1..n:d[i,j]<M}x[i,j];
```

文件 amplEx6.3-6b.txt 中有完整的模型.

习题 6.3E

1. 修改文件 solverEx6.3-6.xls, 求解下面给出节点对之间的最短路径:

   (a) 节点 1 到节点 5.　　(b) 节点 4 到节点 3.

2. 修改 amplEx6.3-6b. txt, 求解习题 6.3A 第 2 题中节点 1 与节点 7 之间的最短路径. 输入的数据是原始的概率. 应用 AMPL 编程技巧来展示这条最优的传输路径以及它传输成功的概率.

6.4 最大流模型

考虑从油井到提炼厂输送原油的管道网络. 为了在网络中顺畅地输送原油, 需要在中间合适的距离安装增压泵站, 每一段管道有一个有限的最大原油流量 (或者称为容量), 每段管道可以是单向的, 也可以是双向的, 这取决于它原始的设计. 图 6.27 描述了一个典型的管道网络. 那么如何确定从油井到提炼厂原油的最大流量?

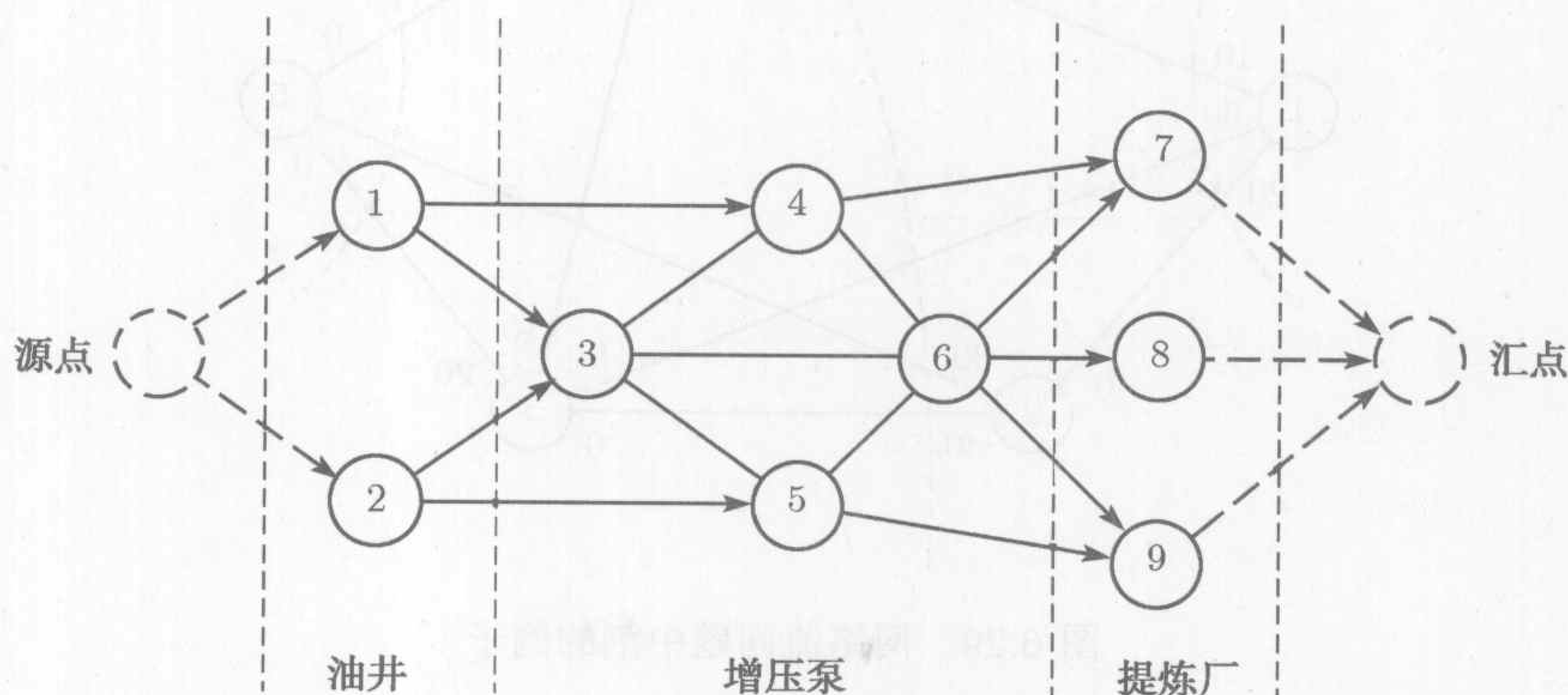


图 6.27　通过增压泵站连接油井与提炼厂、有容量限制的网络

求解上面这个问题时需要添加一个源点, 它连接到每一个油井, 而且是单向的, 每条弧上的容量为无穷大, 还需要添加一个汇点, 每一个提炼厂都连接到汇点, 这些弧上的容量也是无穷大, 如图 6.27 中虚线表示的部分.

给定弧 (i,j), 其中 $i<j$, 我们用符号 $(\overline{C}_{ij},\overline{C}_{ji})$ 表示两个方向 $i\to j$ 和 $j\to i$ 弧上的容量. 为了更清晰地描述, 我们把 $\overline{C}_{ij}$ 放在靠近节点 i 的一边, 把 $\overline{C}_{ji}$ 放在靠近节点 j 的一边, 如图 6.28 所示.

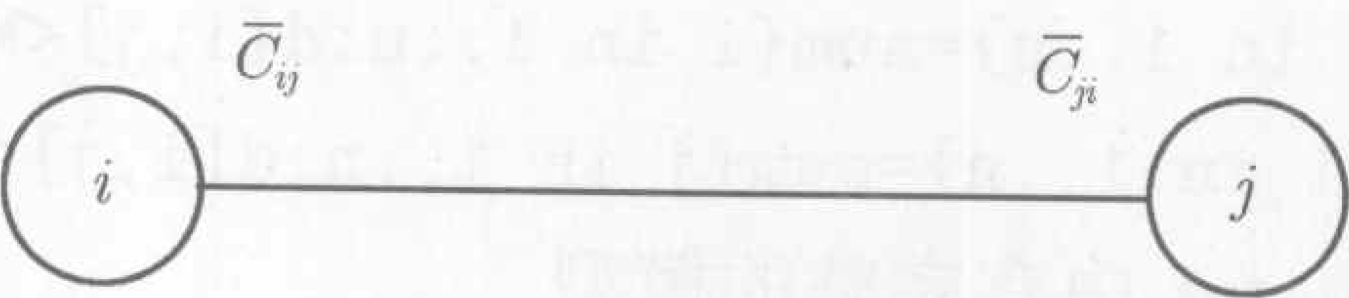


图 6.28　$\overline{C}_{ij}$ 表示从 $i\to j$ 的弧上的容量, $\overline{C}_{ji}$ 表示从 $j\to i$ 的弧上的容量

6.4.1　枚举割

割 (cut) 是一部分弧的集合, 如果将这些弧从网络中删除, 就会断开源点到汇点之间所有的流. **割的容量** (cut capacity) 等于这个割中所有弧上容量的和. 在网络中所有可能的割中, 最小容量的割就对应了这个网络的最大流.

例 6.4-1

考虑图 6.29 中的网路, 按照图 6.28 的表示方法对每一条双向弧都标上了容量. 例如, 对于弧 (3,4), 从节点 3 到节点 4 的流量限制为 10 个单位, 从节点 4 到节点 3 为 5 个单位.

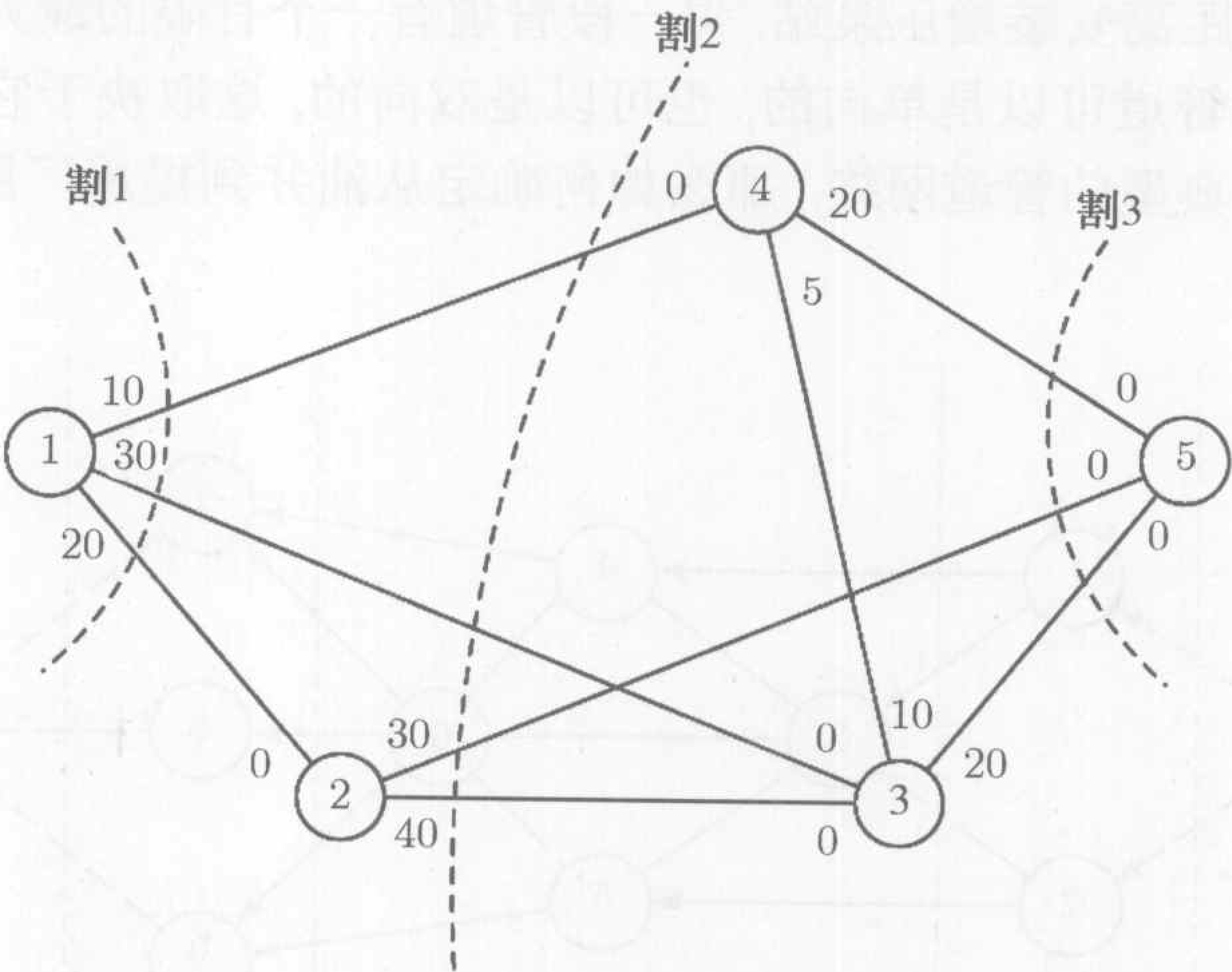


图 6.29　网络流问题中割的例子

图 6.29 中标出的 3 个割的容量见下表.

割	相应的弧	容量
1	(1,2),(1,3),(1,4)	20+30+10=60
2	(1,3),(1,4),(2,3),(2,5)	30+10+40+30=110
3	(2,5),(3,5),(4,5)	30+20+20=70

从这 3 个割上可以得到的信息是, 这个网络的最大流不会超过 60 个单位. 如果要确定最大流, 就必须枚举所有的割, 对一般的网络这显然是一件非常困难的事情. 因此, 寻找有效的算法势在必行.

习题 6.4A

*1. 对于图 6.29 的网络, 确定另外两个割并求出它们的容量.

6.4.2 最大流算法

最大流算法的核心思想就是, 在网络中寻找从源点到汇点具有正的流的**突破路径**. 网络中总的流量就是由这样的一些路径中的流量组成的, 每条路径上的流量是路径上弧的容量或者容量的一部分.

考虑弧 (i,j) 上的初始容量为 $(\overline{C}_{ij},\overline{C}_{ji})$. 当给定了一个流之后, 需要对每条弧上的容量进行修改, 修改后的容量称为**剩余容量**, 用 (c_{ij},c_{ji}) 表示.

如果节点 j 从节点 i 接收到一个流, 那么给节点 j 标号为 $[a_j,i]$, 其中 a_j 表示从节点 i 流到节点 j 的量. 整个算法的步骤如下:

第 1 步 首先对所有的弧 (i,j) 设置剩余容量为初始的容量, 即 $(c_{ij},c_{ji})=(\overline{C}_{ij},\overline{C}_{ji})$. 取 $a_1=\infty$, 给源点 1 标号 $[\infty,—]$. 令 $i=1$, 进入第 2 步.

第 2 步 确定集合 S_i, 它是从节点 i 通过剩余容量为正的弧能够直接到达的所有未标号的节点 j 的集合 (即, 对任意的 $j\in S_i$, 都有 $c_{ij}>0$). 如果 $S_i\neq\varnothing$, 进入第 3 步; 否则, 进入第 4 步.

第 3 步 确定 $k\in S_i$, 使之满足

$$c_{ik}=\max_{j\in S_i}\{c_{ij}\}$$

令 $a_k=c_{ik}$, 并且给节点 k 标号 $[a_k,i]$. 如果 $k=n$, 此时汇点得到了标号, 也就找到了一条*突破路径*, 进入第 5 步; 否则, 令 $i=k$, 转回第 2 步.

第 4 步(回溯) 如果 $i=1$, 那么不存在突破路径, 进入第 6 步; 否则, 令 r 是紧接在当前节点 i 之前得到标号的节点, 并将节点 i 从节点 r 的相邻节点集合中删除, 令 $i=r$, 然后转回第 2 步.

第 5 步(剩余容量的确定) 令 $N_p=(1,k_1,k_2,\cdots,n)$ 是从源点 1 到汇点 n 的第 p 条突破路径. 那么这条路径上最大流量为

$$f_p=\min\{a_1,a_{k_1},a_{k_2},\cdots,a_n\}$$

这条路径上每条弧的剩余容量在顺着流的方向上 f_p 是*减少*的, 在逆着流的方向上 f_p 是*增加*的, 例如, 对于该路径上的节点 i 与节点 j, 当前剩余容量 (c_{ij},c_{ji}) 变为

(a) 如果流是从节点 i 流向节点 j, 那么弧上的剩余容量是 $(c_{ij}-f_p,c_{ji}+f_p)$;

(b) 如果流是从节点 j 流向节点 i, 那么弧上的剩余容量是 $(c_{ij}+f_p, c_{ji}-f_p)$. 恢复在第 4 步删除掉的节点, 然后令 $i=1$, 转回第 2 步接着寻找新的突破路径.

第 6 步(最优解)

(a) 给出已经找到的 m 条突破路径, 那么网络的最大流是

$$F = f_1 + f_2 + \cdots + f_m$$

(b) 根据每条弧 (i,j) 上的*初始的*和*最终的*剩余容量, $(\overline{C}_{ij}, \overline{C}_{ji})$ 和 (c_{ij}, c_{ji}), 按照下面的方法确定这条弧上的最优流: 令 $(\alpha,\beta)=(\overline{C}_{ij}-c_{ij}, \overline{C}_{ji}-c_{ji})$, 如果 $\alpha>0$, 那么从节点 i 流向节点 j 的最优流是 α 个单位; 否则, 如果 $\beta>0$, 那么从节点 j 流向节点 i 的最优流是 β 个单位. (值得注意的是 α 和 β 不可能同时为正数.)

算法第 4 步中的回溯操作是用于解决算法运行中出现中间节点称为"死结"的情形. 算法第 5 步中的流量调整可以通过图 6.30 中一个简单的网络流的例子加以说明. (a) 中网络给出了第一条突破路径 $N_1=\{1,2,3,4\}$, 这条路上的最大流量是 $f_1=5$. 按照算法的第 5 步, 路上每一条弧 $(1,2),(2,3),(3,4)$ 的剩余容量从 $(5,0)$ 变为 $(0,5)$. (b) 中网络给出了第二条突破路径 $N_2=\{1,3,2,4\}$, 这条路上的最大流量为 $f_2=5$. 经过必要的流量调整后得到了 (c) 网络, 此时不存在其他的突破路径. 在 (b) 网络到 (c) 网络所做的工作就是将先前一步从节点 2 到节点 3 的流取消, 算法之所以能做到这一点正是由于在前一步的时候将节点 2 到节点 3 逆向弧的容量从 0 增加到 5(算法第 5 步).

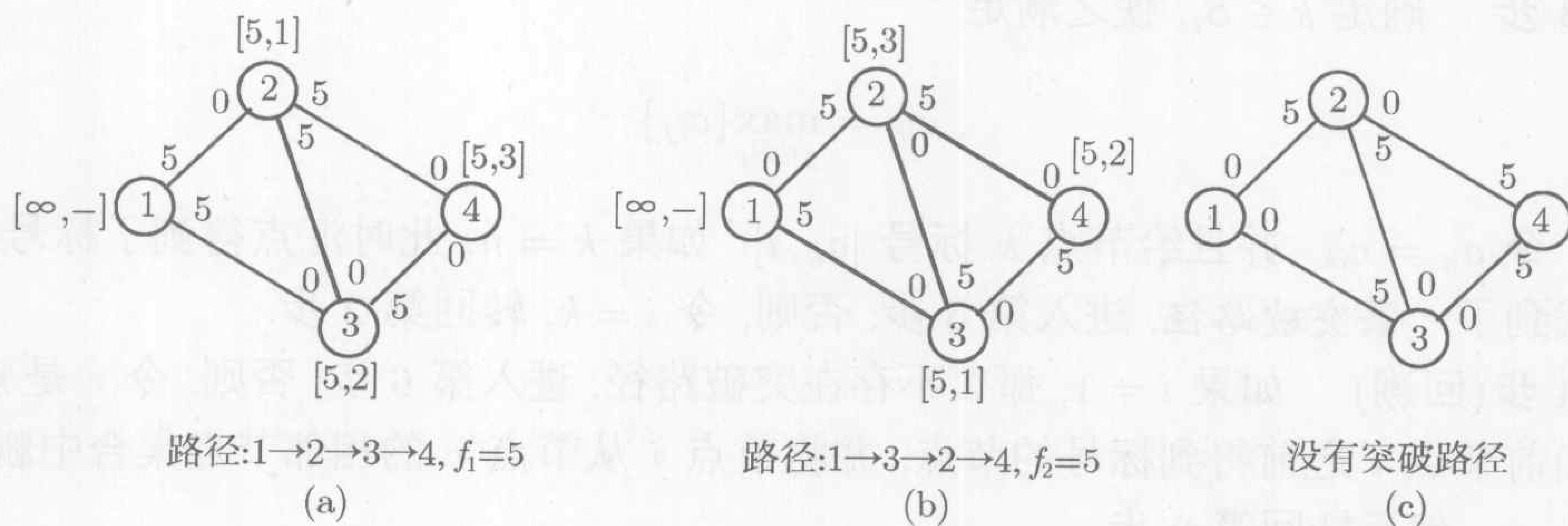


图 6.30 利用剩余容量计算最大流

例 6.4-2

求例 6.4-1(图 6.29) 中网络上的最大流. 图 6.31 形象地给出了算法每一步执行的情况, 这对于理解算法会有很大帮助.

迭代 1 令初始的剩余容量 (c_{ij}, c_{ji}) 等于初始容量 $(\overline{C}_{ij}, \overline{C}_{ji})$.

第 1 步 取 $a_1=\infty$, 给源点 1 标号 $[\infty,-]$. 令 $i=1$.

第 2 步 $S_1=\{2,3,4\}(\neq\varnothing)$.

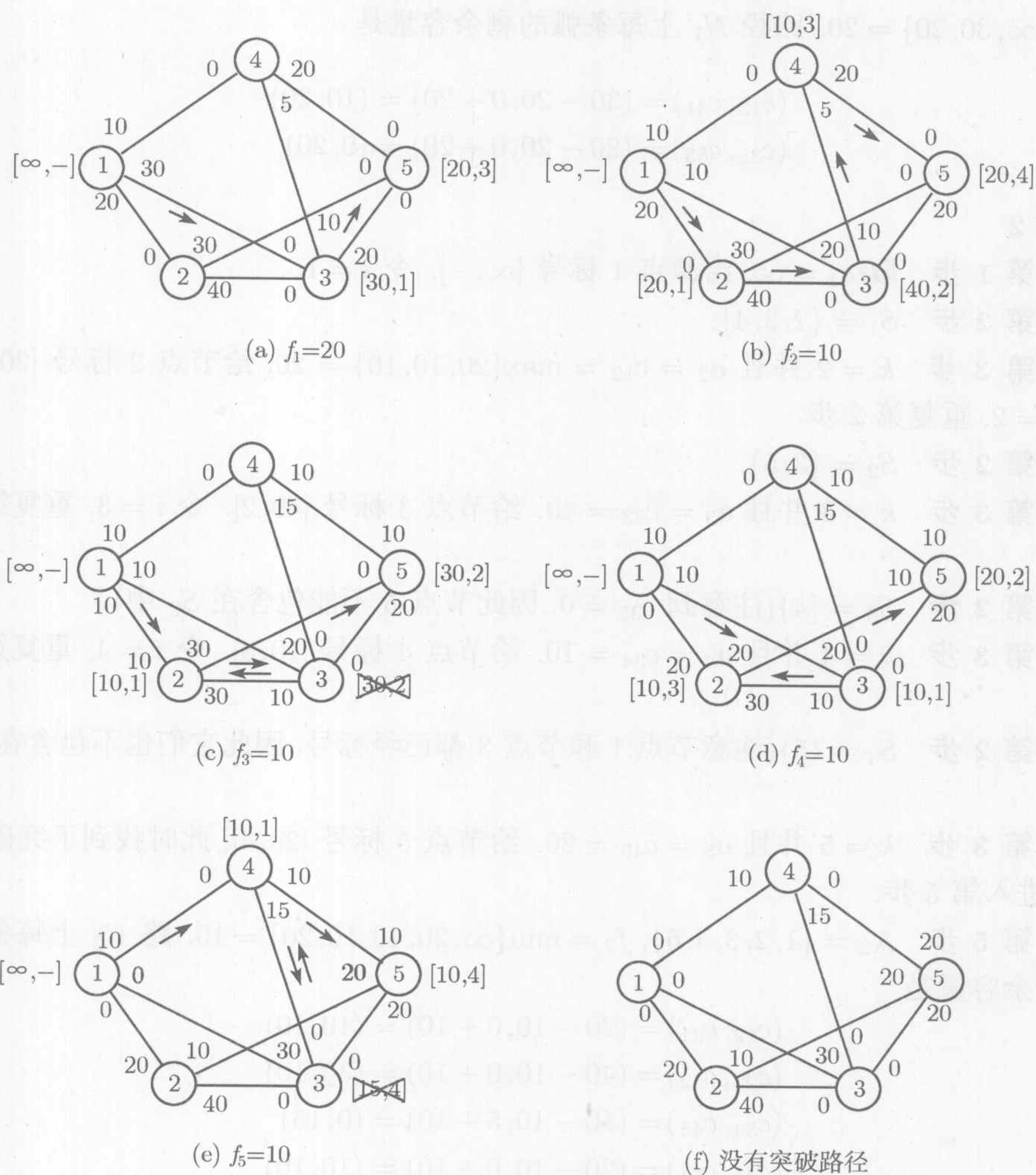


图 6.31 例 6.4-2 中最大流算法的迭代过程

第 3 步 $k = 3$, 因为 $c_{13} = \max\{c_{12}, c_{13}, c_{14}\} = \max\{20, 30, 10\} = 30$. 令 $a_3 = c_{13} = 30$, 给节点 3 标号 [30, 1]. 令 $i = 3$, 重复第 2 步.

第 2 步 $S_3 = \{4, 5\}$.

第 3 步 $k = 5$ 并且 $a_5 = c_{35} = \max\{10, 20\} = 20$. 给节点 5 标号 [20, 3], 此时找到了突破路径, 进入第 5 步.

第 5 步 根据从节点 5 向后直到节点 1 的标号确定出这条突破路径, 即 $(5) \rightarrow [20, 3] \rightarrow (3) \rightarrow [30, 1] \rightarrow (1)$. 那么, $N_1 = \{1, 3, 5\}$, $f_1 = \min\{a_1, a_3, a_5\} =$

$\min\{\infty,30,20\}=20$. 路径 N_1 上每条弧的剩余容量是

$$(c_{13},c_{31})=(30-20,0+20)=(10,20)$$
$$(c_{35},c_{53})=(20-20,0+20)=(0,20)$$

迭代 2

第 1 步　取 $a_1=\infty$, 给源点 1 标号 $[\infty,-]$. 令 $i=1$.

第 2 步　$S_1=\{2,3,4\}$.

第 3 步　$k=2$ 并且 $a_2=c_{12}=\max\{20,10,10\}=20$, 给节点 2 标号 $[20,1]$. 令 $i=2$, 重复第 2 步.

第 2 步　$S_2=\{3,5\}$.

第 3 步　$k=3$ 并且 $a_3=c_{23}=40$. 给节点 3 标号 $[40,2]$. 令 $i=3$, 重复第 2 步.

第 2 步　$S_3=\{4\}$(注意到 $c_{35}=0$, 因此节点 5 不能包含在 S_3 中).

第 3 步　$k=4$ 并且 $a_4=c_{34}=10$. 给节点 4 标号 $[10,3]$. 令 $i=4$, 重复第 2 步.

第 2 步　$S_4=\{5\}$(注意节点 1 和节点 3 都已经标号, 因此它们也不包含在 S_4 中).

第 3 步　$k=5$ 并且 $a_5=c_{45}=20$. 给节点 5 标号 $[20,4]$, 此时找到了突破路径, 进入第 5 步.

第 5 步　$N_2=\{1,2,3,4,5\}$, $f_2=\min\{\infty,20,40,10,20\}=10$. 路 N_2 上每条弧的剩余容量是

$$(c_{12},c_{21})=(20-10,0+10)=(10,10)$$
$$(c_{23},c_{32})=(40-10,0+10)=(30,10)$$
$$(c_{34},c_{43})=(10-10,5+10)=(0,15)$$
$$(c_{45},c_{54})=(20-10,0+10)=(10,10)$$

迭代 3

第 1 步　取 $a_1=\infty$, 给源点 1 标号 $[\infty,-]$. 令 $i=1$.

第 2 步　$S_1=\{2,3,4\}$.

第 3 步　$k=2$ 并且 $a_2=c_{12}=\max\{10,10,10\}=10$(虽然可以任意选择一个, TORA 选择序号最小的节点. 以下也使用同样的方法.), 给节点 2 标号 $[10,1]$. 令 $i=2$, 重复第 2 步.

第 2 步　$S_2=\{3,5\}$.

第 3 步　$k=3$ 并且 $a_3=c_{23}=30$. 给节点 3 标号 $[30,2]$. 令 $i=3$, 重复第 2 步.

第 2 步　$S_3=\varnothing$(因为 $c_{34}=c_{35}=0$). 进入第 4 步回溯.

第 4 步回溯 从节点 3 的标号 [30, 2] 可以立即知道前一个节点是 $r=2$, 在这一次的迭代中将节点 3 删除不再考虑. 令 $i=r=2$, 重复第 2 步.

第 2 步 $S_2=\{5\}$(注意节点 3 已经在回溯中删除).

第 3 步 $k=5$ 并且 $a_5=c_{25}=30$. 给节点 5 标号 [30, 2]. 此时找到了一条突破路径, 进入第 5 步.

第 5 步 $N_3=\{1,2,5\}$, $f_3=\min\{\infty,10,30\}=10$. 路 N_3 上每条弧的剩余容量是

$$(c_{12},c_{21})=(10-10,10+10)=(0,20)$$
$$(c_{25},c_{52})=(30-10,0+10)=(20,10)$$

迭代 4 这一步得到的路是 $N_4=\{1,3,2,5\}$, $f_4=10$. (请验证!)

迭代 5 这一步得到的路是 $N_5=\{1,4,5\}$, $f_5=10$. (请验证!)

迭代 6 此时所有从节点 1 出去的弧的剩余容量都是 0, 因此不再存在其他的突破路径. 转入第 6 步给出最优解.

第 6 步 网络的最大流是 $F=f_1+f_2+\cdots+f_5=20+10+10+10+10=60$ 个单位. 每条弧上对应的流量是用初始的容量 $(\overline{C}_{ij},\overline{C}_{ji})$ 减掉迭代 6 最后得到的剩余容量 (c_{ij},c_{ji}), 见下表:

弧	$(\overline{C}_{ij},\overline{C}_{ji})-(c_{ij},c_{ji})_6$	流量	方向
(1,2)	$(20,0)-(0,20)=(20,-20)$	20	$1\to2$
(1,3)	$(30,0)-(0,30)=(30,-30)$	30	$1\to3$
(1,4)	$(10,0)-(0,10)=(10,-10)$	10	$1\to4$
(2,3)	$(40,0)-(40,0)=(0,0)$	0	—
(2,5)	$(30,0)-(10,20)=(20,-20)$	20	$2\to5$
(3,4)	$(10,5)-(0,15)=(10,-10)$	10	$3\to4$
(3,5)	$(20,0)-(0,20)=(20,-20)$	20	$3\to5$
(4,5)	$(20,0)-(0,20)=(20,-20)$	20	$4\to5$

TORA 程序

利用 TORA 来求解最大流模型, 既可以按照自动模式求解, 也可以按照上面的迭代一步步给出详细的过程. 从菜单 SOLVER/MODITY 中选择 Solve Problem. 在指明了输出格式之后, 到输出窗口中选择 Maximum Flows 或者 Iterations. 文件 toraEx6.4-2.txt 提供了求解例 6.4-2 的 TORA 数据.

习题 6.4B

*1. 在例 6.4-2 中:

(a) 确定每条弧上的剩余容量.

(b) 确定通过节点 2, 3, 4 的总流量.

(c) 如果增加 $3 \to 5$ 和 $4 \to 5$ 上的容量, 那么网络的最大流是否可以增加?

2. 求图 6.32 中网络上的最大流以及对应最大流的每条弧上的流量.

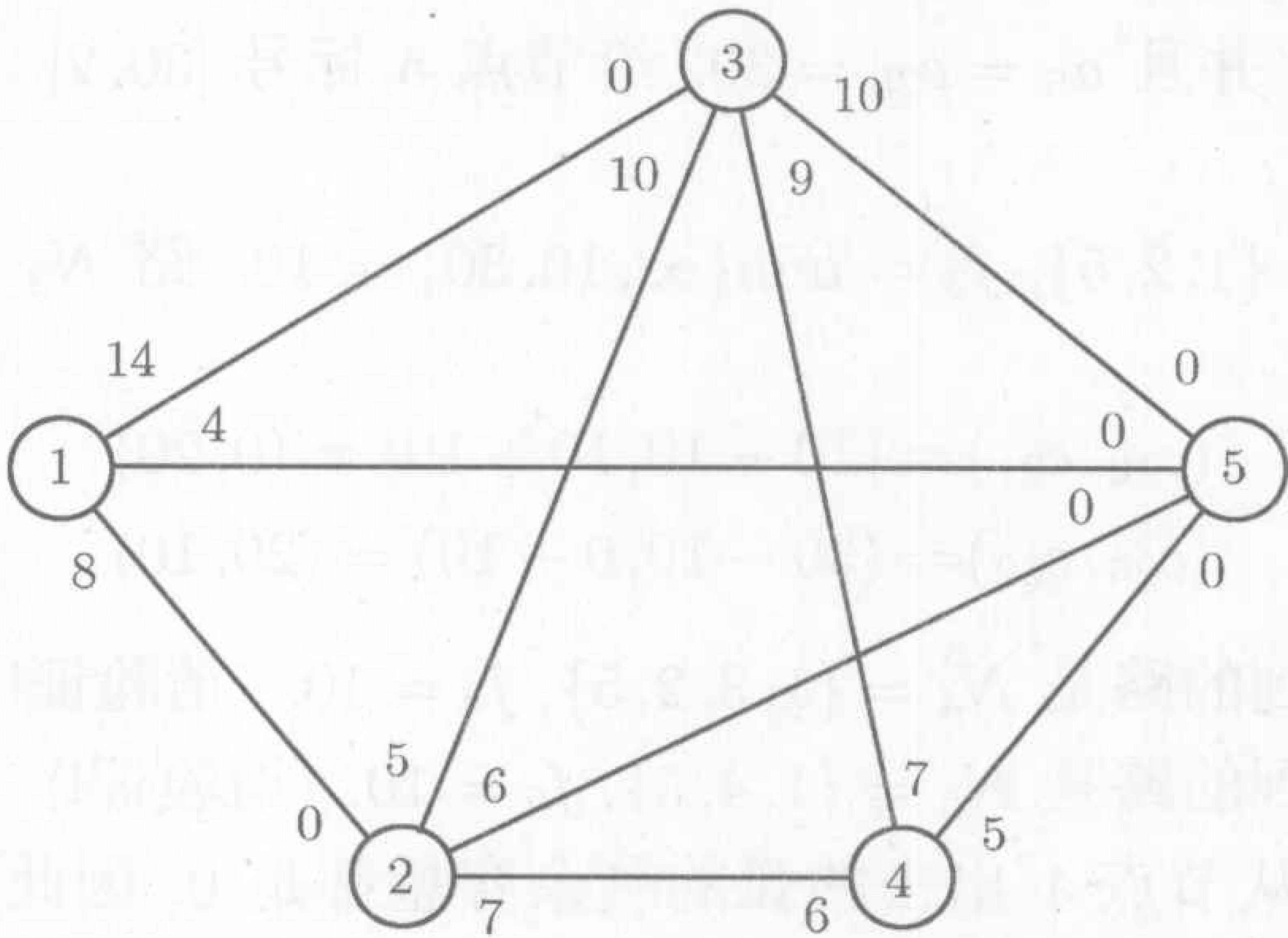


图 6.32 习题 6.4B 第 2 题中的网络

3. 3 个炼油厂通过管道网络为 2 个分销终端运送汽油. 如果这个网络运送能力满足不了需求, 那么可以从其他的源点获取. 这个管道网络中有 3 个泵站, 如图 6.33, 汽油的流向如图中箭头所示. 图中标出了每一段管道的运送容量, 单位是 100 万桶/天. 求解下面的问题:

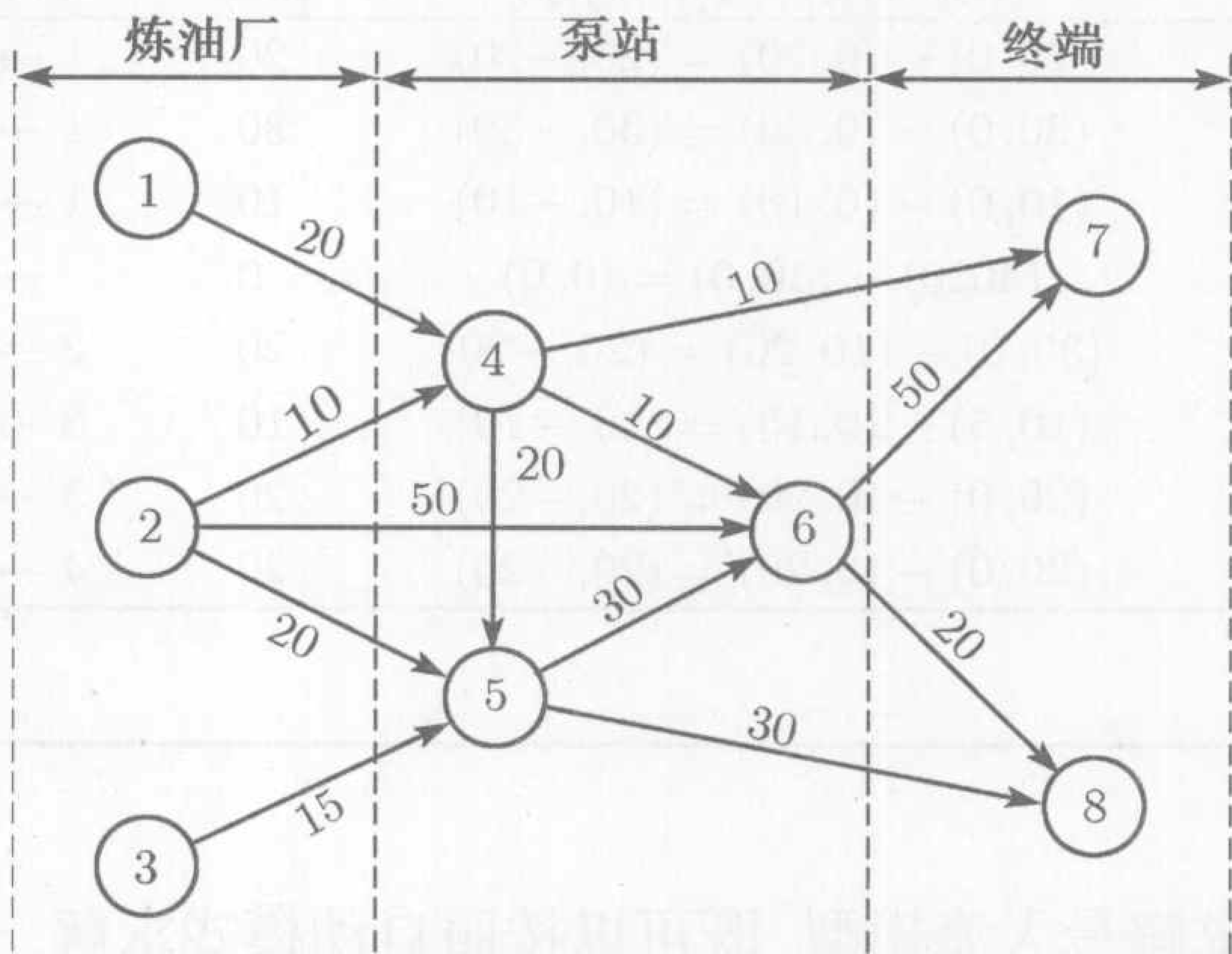


图 6.33 习题 6.4B 第 3 题中的网络

(a) 要满足这个网络的最大流量, 那么每一个炼油厂每天的产量应该是多少.

(b) 要满足这个网络的最大流量, 那么每一个终端每天的需求量应该是多少.

(c) 要满足这个网络的最大流量, 每个泵站的每天容量应该是多少.

4. 假定在图 6.33 网络中泵站 6 的每天最大容量限制为 5 000 万桶. 根据这个约束重新建立网络模型, 并求出相应的网络的最大容量.

5. 需要将鸡饲料从 3 个存放仓库用卡车运送到 4 个农场, 某些仓库的饲料不能够直接运送到某些农场. 一条线路上的容量取决于这条路上可用的卡车数目以及卡车每天往返的次数.

下表给出了仓库每天可以供应的饲料总数以及农场每天的需求量 (单位是 1 000 磅), 另外表中单元格内给出了相应路线上的每天运送的最大容量.

		农场 1	农场 2	农场 3	农场 4	
仓库	1	30	5	0	40	**20**
	2	0	0	5	90	**20**
	3	100	40	30	40	**200**
		200	**10**	**60**	**20**	

(a) 求满足最大需求量的运送计划.

(b) 这个运送计划是否能够满足所有农场的需求?

6. 在第 5 题中, 假定同时也允许仓库 1 与仓库 2 以及仓库 2 与仓库 3 之间运送饲料. 还假定农场 1 与农场 2, 农场 2 与农场 3, 农场 3 与农场 4 之间也可以相互运送饲料, 这些路线上每天的运送容量是 5 万磅. 这样对于不满足需求的农场之间的运送有什么影响?

***7.** 一对夫妇有 5 个孩子, 现在有 5 项家务需要分配给他们. 根据以往的经验, 如果将家务强加于一个孩子, 那么他就因反感而做不完. 考虑到这一点, 父母会询问每个孩子, 各自喜欢做哪一项家务, 结果如下表:

孩子	喜欢做的家务
Rif	3, 4, 5
Mai	1
Ben	1, 2
Kim	1, 2, 5
Ken	2

父母需要做的就是尽可能多地将家务分配给喜欢做的孩子. 求最好的分配方案使得可以完成的家务数目最大, 并求出这个数目.

8. 有 4 家工厂从事于生产 4 种玩具. 下表给出了每家工厂可以生产的玩具种类.

工厂	可以生产的玩具种类
1	1, 2, 3
2	2, 3
3	1, 4
4	3, 4

所有种类的玩具生产一件需要的工时和原料基本相同. 4 家工厂每天的产量分别是 200, 180, 300, 100 个玩具. 对 4 种玩具的需求量分别是 200, 150, 350 和 100 个. 求工厂可以最大满足对 4 种玩具需求的生产方案.

9. A 大学的学术委员会需要从来自 4 个社团的 6 名学生中选择代表, 学术委员会的代表要熟悉下面 3 个领域: 数学、艺术和工程. 每个领域最多有 2 名学生进入委员会. 下表给出了这 6 名学生隶属于哪些社团的情况:

社团	隶属的学生
1	1, 2, 3
2	1, 3, 5
3	3, 4, 5
4	1, 2, 4, 6

这 6 名学生所擅长的领域列表如下：

领域	擅长的学生
数学	1, 2, 4
艺术	3, 4
工程	4, 5, 6

即便一名学生擅长于多个领域, 他也只能被分配到一个特定的领域. 请问委员会是否能够代表所有 4 个社团?

10. *网络中弧上有流量下界要求的最大/最小流.* 在刚刚这一节中介绍的最大流算法是针对网络中弧上的流量下界要求是 0. 然而在某些模型中, 网络中弧上的流量下界要求是严格正的, 在这种情况下我们感兴趣的是找到网络中的最大或者最小流 (参见附录 E 中的案例 6-3). 弧上有了流量下界要求之后会带来许多困难, 因为这种情况下的网络可能根本就没有可行流. 本题的目的就是要证明, 对于弧上有流量下界为正数要求的最大流和最小流模型, 可以通过下面的两个步骤来求解.

第 1 步 对于弧上有流量下界为正数要求的网络, 首先找到一个初始的可行解.

第 2 步 根据第一步中得到的可行解, 寻找原始网络中的最大或者最小流.

(a) 证明对于一条有流量限制为 $l_{ij} \leqslant x_{ij} \leqslant u_{ij}$ 的弧 (i, j), 可以等价于用下面的方法表示：在节点 i 处增加一个需求量为 l_{ij} 的*汇点*, 以及在节点 j 处增加一个供应量为 l_{ij} 的*源点*, 然后弧 (i, j) 上的流量限制为 $0 \leqslant x_{ij} \leqslant u_{ij} - l_{ij}$.

(b) 证明寻找原始网络上的一个可行流等价于寻找经过下面转化之后的网络上的最大流 x'_{ij}：(1) 将流量 x'_{ij} 的上下界改为 $0 \leqslant x'_{ij} \leqslant u_{ij} - l_{ij}$; (2) 将所有增加的源点连接到一个超级源点并且出弧 (方向从超级源点指向增加的源点) 的容量为 l_{ij}; (3) 将所有增加的汇点连接到一个超级汇点并且入弧 (方向指向超级汇点) 的容量为 l_{ij}; (4) 在原始网络中连接终端 t 到源点 s 并且弧上的容量为无穷. 如果新网络中的最大流等于原始网络中的全部下界之和, 则可行解存在. 将上面的过程应用到下面给出的网络中, 并求出它的一个可行解.

(c) 根据 (b) 得到的可行解, 应用最大流算法求原始网络上的*最小流*. (*提示*：首先根据给出的初始可行解计算出剩余容量网络, 然后求从*终点到始发点*的最大流, 这样做等价于确定从始发点到终点应当取消的最大流. 最后结合可行解与这个最大流就可以得到原始网络中的最小流.)

(d) 根据 (b) 得到的可行解, 应用最大流算法求原始网络上的最大流. [提示: 与 (c) 中类似, 首先从剩余容量网络出发, 然后以得到的可行解为初始流在这个网络上应用最大流算法.]

弧 (i,j)	(l_{ij},u_{ij})
(1,2)	(5,20)
(1,3)	(0,15)
(2,3)	(4,10)
(2,4)	(3,15)
(3,4)	(0,20)

6.4.3 最大流问题的线性规划模型

定义 x_{ij} 为弧 (i,j) 上的流量, 其中 (i,j) 的容量为 C_{ij}. 目标就是确定所有的 x_{ij} 使得从始发点 s 到终点 t 的流量最大, 同时满足在除了 s 和 t 之外的所有节点上的流守恒 (输入流 = 输出流).

例 6.4-3

在图 6.29 的最大流模型中 (例 6.4-2), $s=1, t=5$. 下表给出相应的线性规划模型, 模型有两个不同但是却等价的目标函数, 一个是最大化始发点 $1(=z_1)$ 的输出流, 另一个是最大化终点 $5(=z_2)$ 的输入流.

	x_{12}	x_{13}	x_{14}	x_{23}	x_{25}	x_{34}	x_{35}	x_{43}	x_{45}	
max $z_1=$	1	1	1							
max $z_2=$					1		1		1	
节点 2	1			−1	−1					= 0
节点 3		1		1		−1	−1	1		= 0
节点 4			1			1		−1	−1	= 0
容量	20	30	10	40	30	10	20	5	20	

利用任何一个目标求出的最优解都是

$$x_{12}=20, x_{13}=30, x_{14}=10, x_{25}=20, x_{34}=10, x_{35}=20, x_{45}=20$$

相应的最大流是 $z_1=z_2=60$.

Excel 规划求解程序

图 6.34 给出了针对例 6.4-2 最大流模型的 Excel 规划求解模型(文件 solverEx 6.4-2.xls). 这个模型的基本思想与在例 6.3-6 中详细讲述的最短路径模型类似. 几

点主要的区别是: (1) 在初始点 1 和终点 5 上没有了流方程; (2) 目标是求在节点 1(F9) 上的最大输出流量, 或者求在终点 5(G13) 上的最大输入流量. 文件 solverEx 6.4-2.xls 中将 G13 作为目标单元格. 读者可以练习将 F9 作为目标单元格求解模型.

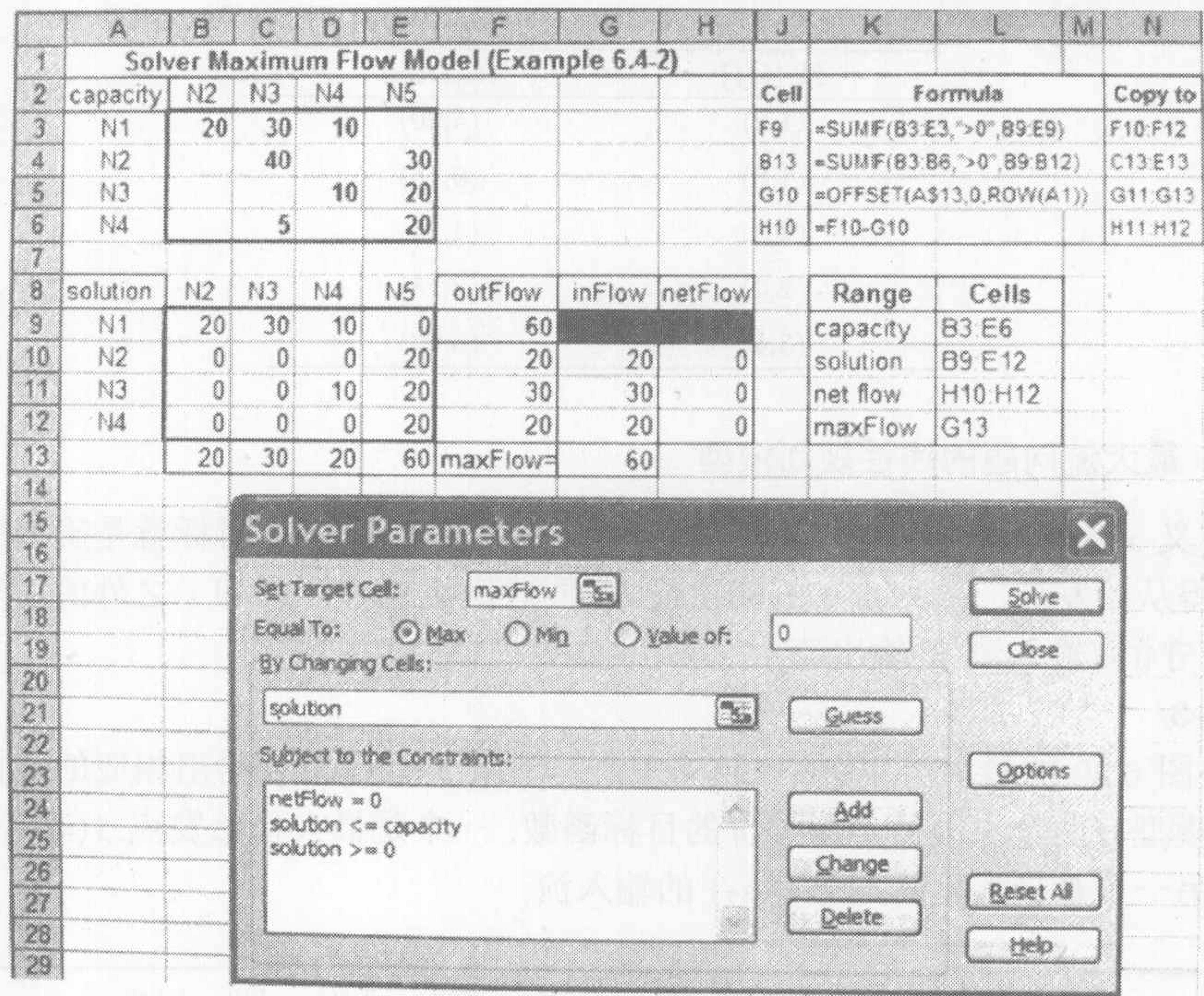


图 6.34　Excel 规划求解关于例 6.4-2 中最大流模型的结果 (文件 solverEx6.4-2.xls)

AMPL 程序

图 6.35 给出了求最大流问题的 AMPL 模型, 使用的是例 6.4-2 中的数据 (文件 amplEx6.4-2.txt). 确定每个节点上输入和输出流的总体思想与求解例 6.3-6 中的最短路径模型非常类似 (所以仔细阅读文件 amplEx6.3-6a.txt 和 amplEx6.3-6b.txt 对理解本问题会有很大帮助). 然而这个模型可以用来求网络中任意两个节点`start`和`end`之间的最大流, 所以需要两个额外约束来保证没有流进入`start`以及没有流离开`end`. 模型中的约束`inStart`和`outEnd`可以保证上面的要求. 当`start=1`、`end=5`时就不需要上面的这两个约束了, 因为这个例题输入数据的性质就可以保证要求的结果. 然而, 如果`start=3`, 由于节点 3 同时允许输入和输出流 (弧 4→3 和弧 3→4), 所以必须要有约束`inStart`(读者可以试着在没有`inStart`约束的情况求解模型!).

目标函数是最大化在节点`start`上总的输出流, 或者也可以最大化在节点`end`上总的输入流. 这个模型可以用于求解网络中任意指定的节点`start` 和`end`之间的最大流.

```
#--------- Maximal Flow model (Example 6.4-2)---------
param n;
param start;
param end;
param c{i in 1..n, j in 1..n} default 0;

var x{i in 1..n,j in 1..n:c[i,j]>0}>=0,<=c[i,j];
var outFlow{i in 1..n}=sum{j in 1..n:c[i,j]>0}x[i,j];
var inFlow{i in 1..n}=sum{j in 1..n:c[j,i]>0}x[j,i];

maximize z: sum {j in 1..n:c[start,j]>0}x[start,j];
subject to
limit{i in 1..n:
      i<>start and i<>end}:outFlow[i]-inFlow[i]=0;
inStart:sum{i in 1..n:c[i,start]>0}x[i,start]=0;
outEnd:sum{j in 1..n:c[end,j]>0}x[end,j]=0;

data;
param n:=5;
param start:=1;
param end:=5;
param c:
  1 2  3  4  5 :=
1 . 20 30 10 0
2 . .  40 0  30
3 . .  0  10 20
4 . .  5  .  20
5 . .  .  .  .;

solve;
print "MaxFlow between nodes",start,"and",end, "=",z;
printf "Associated flows:\n";
for {i in 1..n-1} for {j in 2..n:c[i,j]>0}
    {if x[i,j]>0 then
    printf"(%2i-%2i)= %5.2f\n",i,j,x[i,j];} print;
```

图 6.35 求解例 6.4-2 的最大流问题的 AMPL 模型 (文件 amplEx6.4-2.txt)

习题 6.4C

1. 对下列问题建立线性规划模型, 并利用 Excel 规划求解和 AMPL 求解.

(a) 习题 6.4B 第 2 题. (b)习题 6.4B 第 5 题. (c) 习题 6.4B 第 9 题.

2. Jim 住在科罗拉多州的丹佛, 每年假期都会去怀俄明州的黄石国家公园度假. 由于 Jim 是一位爱好大自然的旅游者, 所以他希望每年开车去度假所选择的线路互不相同. 经过仔细地查看地图之后, Jim 确定了几条从丹佛 (D) 到黄石 (Y) 的路线, 如图 6.36 所示. Jim

希望选择一些从 D 到 Y 的路线, 并且任何两条路线上都没有相同的城市.(利用 AMPL 和 Excel 规划求解), 确定 Jim 可以选择的所有不同的路线. (提示: 修改最大流的线性规划模型, 来求解从 D 到 Y 的两两不同的路径最大数目.)

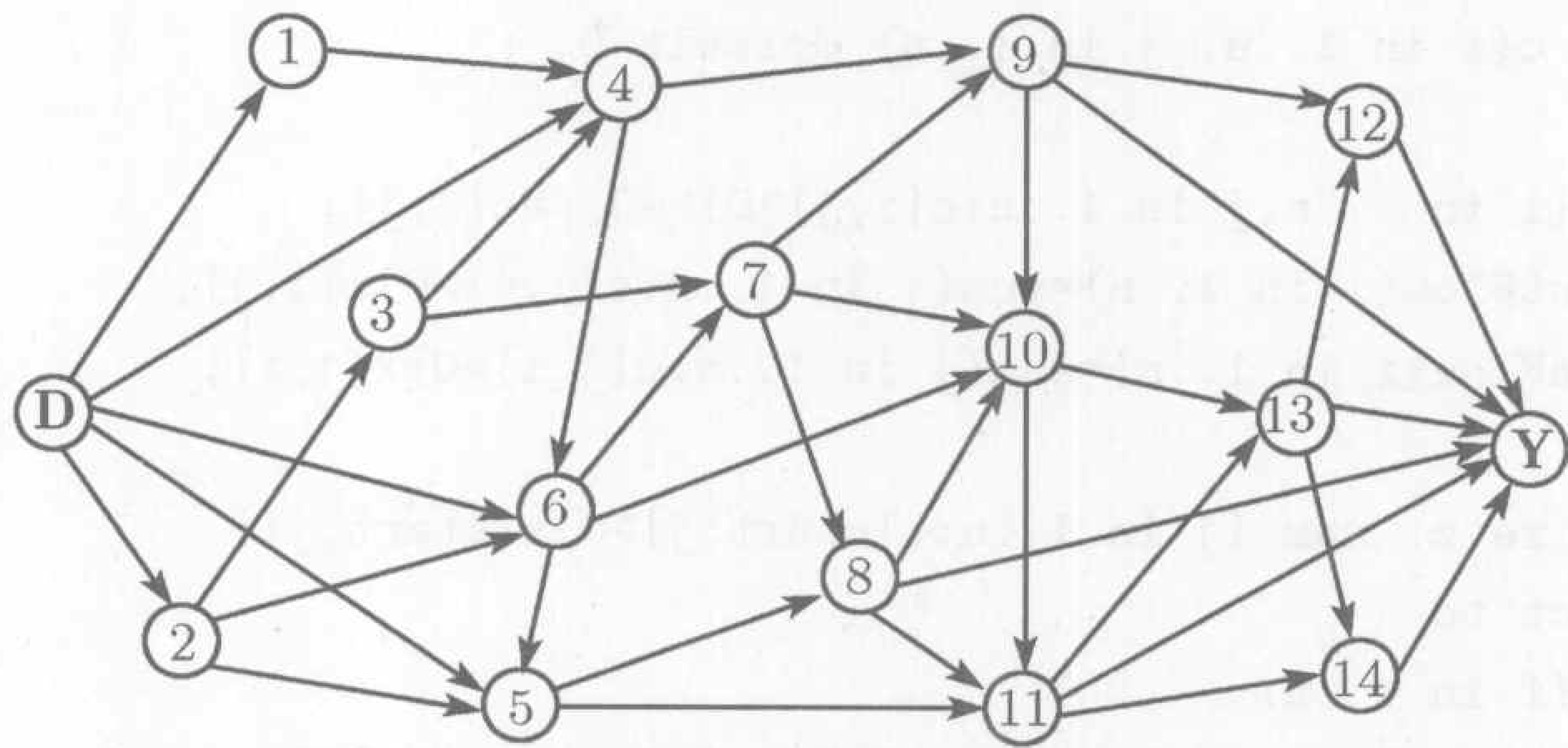


图 6.36　习题 6.4C 第 2 题中的网络

3. (Guéret 等人, 2002, 12.1 节) 一个军用的无线电通信系统由 9 个站点连接组成, 如图 6.37 所示. 即使其他任何 3 个站点都被敌人破坏, 也要能保证站点 4 和站点 7 之间可以相互通信. 请问图中给出的通信网络是否能够满足上面的要求? 利用 AMPL 和 Excel 规划求解来解决这个问题.

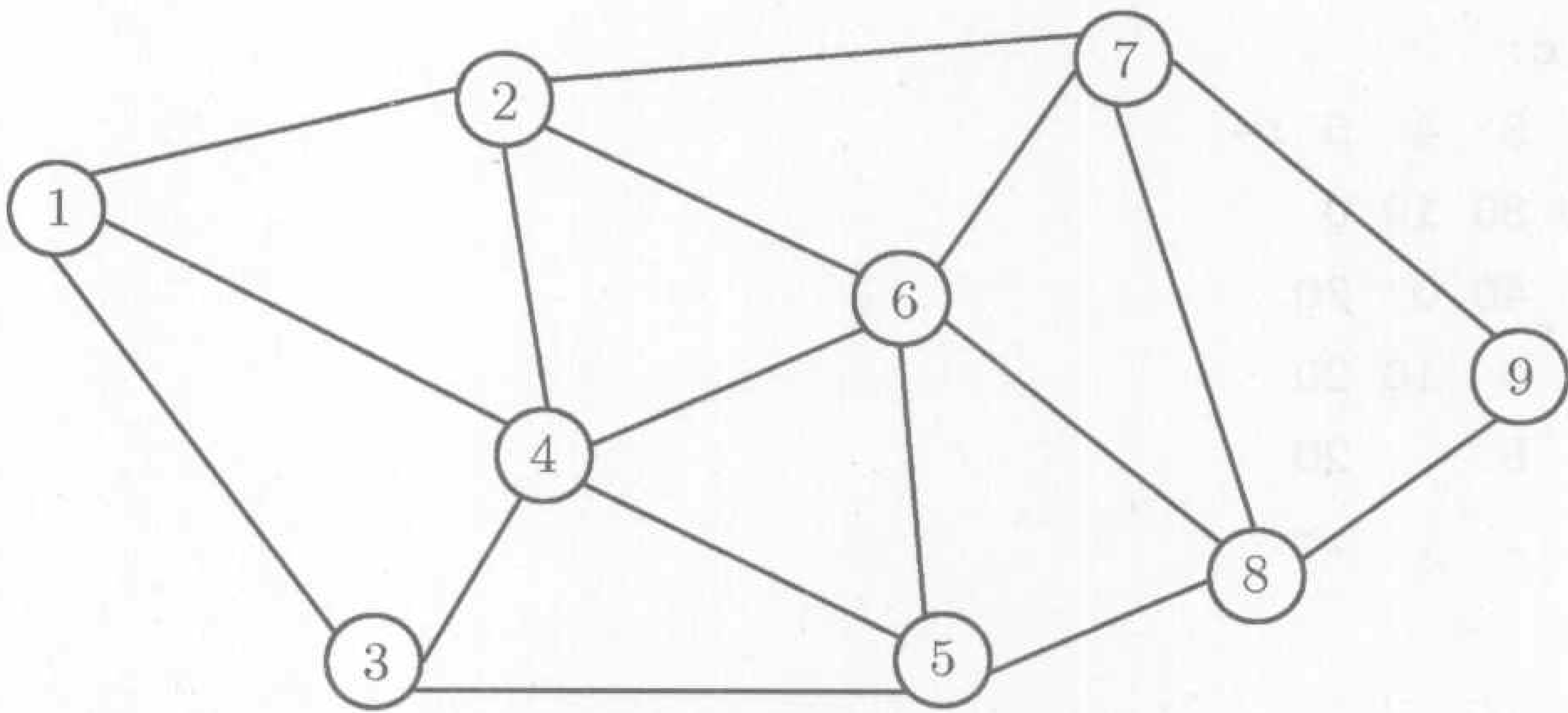


图 6.37　习题 6.4C 第 3 题中的网络

6.5　关键路径方法和计划评审技术

CPM(Critical Path Method, 关键路径方法) 和 PERT(Program Evaluation and Review Technique, 计划评审技术) 是利用网络为基础的方法, 用于帮助制定项目的计划、时间表以及项目的管理. 一个项目是由一组相互关联的活动组成, 其中每项活动都会消耗相应的时间和资源. CPM 和 PERT 的目标就是给这些活动设计一个合理的安排. 图 6.38 给出了这两种方法的基本工作框架. 首先, 需要定义项目中的所有活动、它们之间的先后顺序关系, 以及每个活动的耗时. 其次, 根据活动之间的关系构建网络. 再次, 为该项目计算并规划出相应的时间表. 但是在实际计划执行中, 由于某些活动的提前完成或者延误完成, 使得执行的过程并非完全按照事件

指定的顺序, 在这种情况下, 项目时间表需要修正来与实际相吻合. 这也正是图中为什么要在时间表阶段与构建网络阶段之间包含一个反馈的环, 如图 6.38.

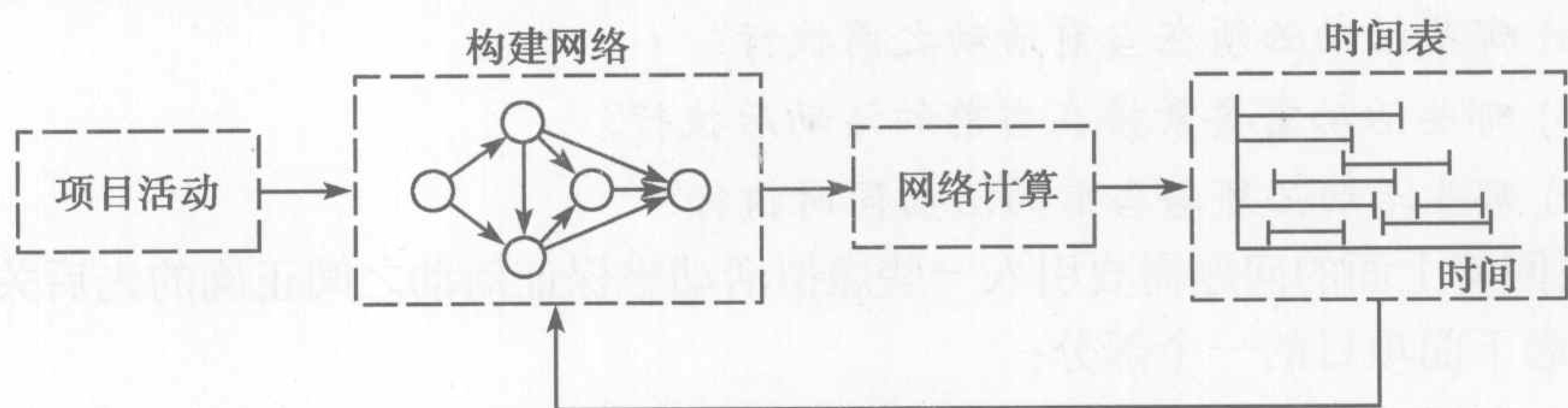


图 6.38 应用 CPM-PERT 指定项目计划的框图

CPM 和 PERT 是两个相互的独立方法, 区别就是 CPM 用于解决活动耗时确定的情形, 而 PERT 是用于耗时不确定的情形. 首先介绍 CPM, 然后再详细介绍 PERT.

6.5.1 网络表示

项目中的每一项活动用一条弧表示, 弧的方向就是项目进行的方向. 网络中的节点表示不同活动之间的先后关系.

在建立网络的时候需要遵守下面 3 条原则.

原则 1 *每一项活动都恰好用一条弧表示.*

原则 2 *每一项活动都必须由两个不同的端点唯一确定.*

图 6.39 给出了如何引入一项虚拟的活动来表示两项同时执行的活动 A 和 B. 定义虚拟活动一般都是用一条虚线表示, 并且虚拟活动消耗的时间和资源都是 0. 为了保证活动 A 和 B 发生的同时性, 需要按照图 6.39 给出的 4 种方法之一来插入一项虚拟活动, 这样做之后, 就可以保证两项活动 A 和 B 都有唯一确定的端点 (满足原则 2).

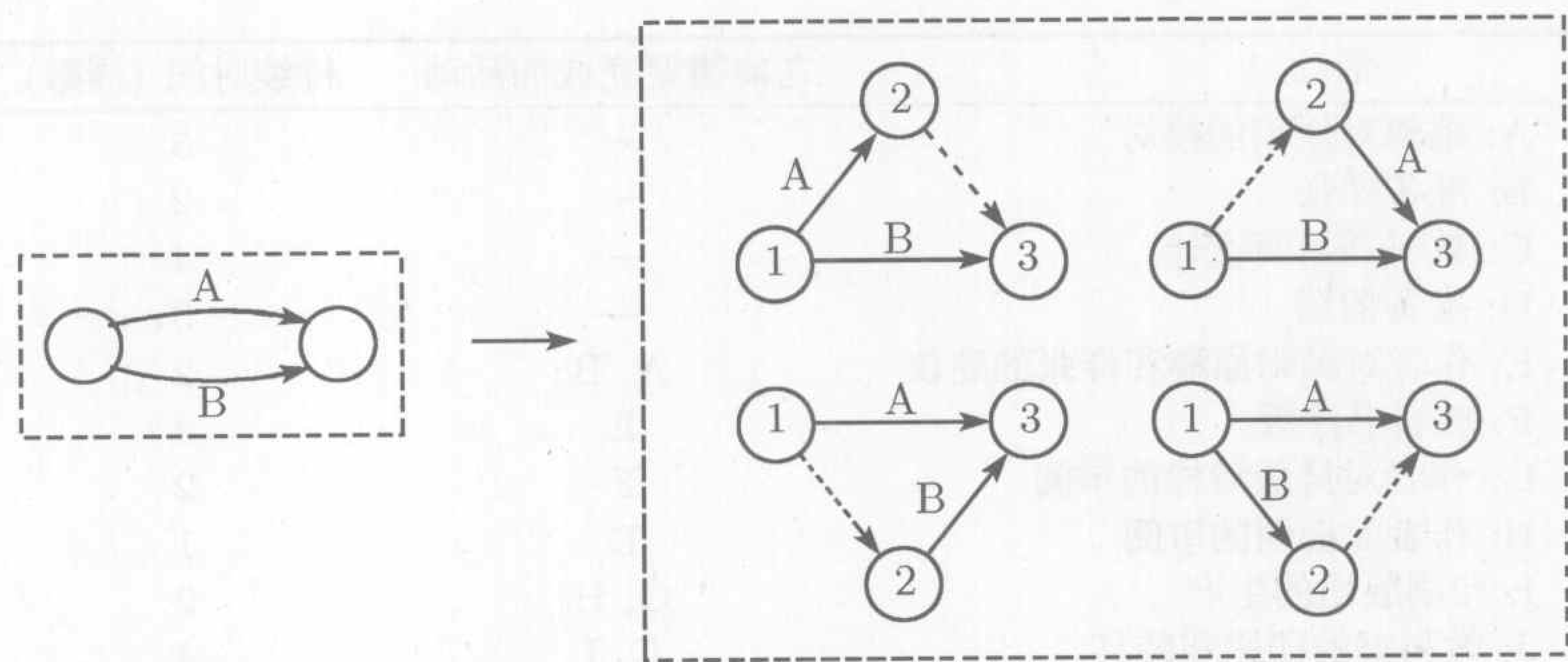


图 6.39 利用虚拟活动来唯一表示两项同时执行的活动

原则 3　为了保证活动之间正确的先后关系, 在将每一项活动加入到网络中时需要回答下面的几个问题:

(a) 哪些活动必须在当前活动之前执行?

(b) 哪些活动需要紧接在当前的活动后执行?

(c) 哪些活动必须与当前的活动同时执行?

要回答上面的问题需要引入一些虚拟活动来保证活动之间正确的先后关系. 例如, 考虑下面项目的一个部分:

(1) 在活动 A 和 B 完成之后立刻执行活动 C.

(2) 活动 E 只需要在活动 B 完成之后就可以执行.

对于这个问题, 图 6.40 中 (a) 给出了不正确的先后关系表示, 因为从这种表示可以看出, 活动 E 也同样是要求在活动 A 和活动 B 都完成之后才可以开始执行. 在 (b) 中, 通过引入一个虚拟的活动 D 解决了这个问题.

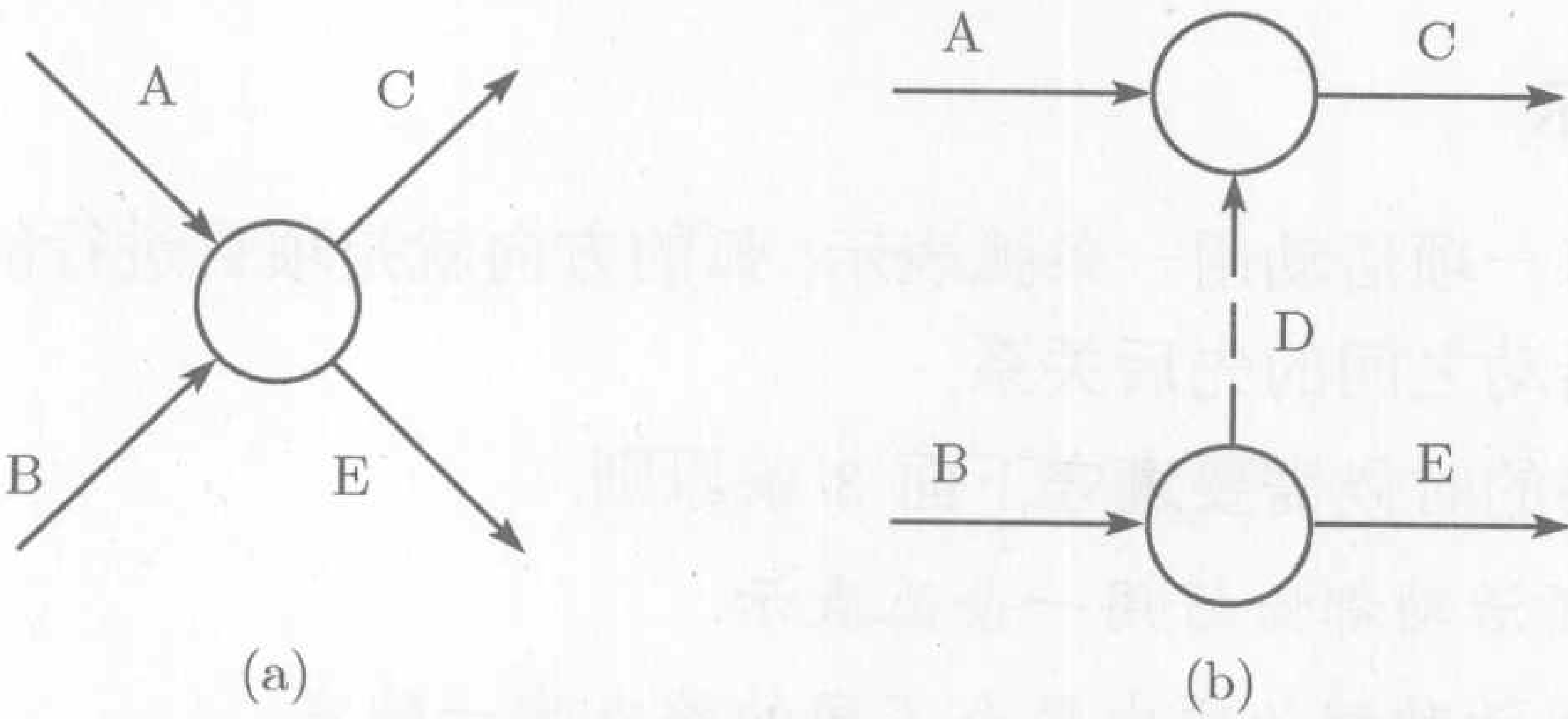


图 6.40　利用虚拟活动来保证活动之间正确的先后关系

例 6.5-1

出版商与一位作者签订了一份出版一本教科书的合同. 下表给出了出版这本教科书的相关 (简化的) 活动. 作者需要向出版商提供教科书原稿的打印版和电子版. 建立这个项目相应的网络.

活动	之前需要完成的活动	持续时间 (周数)
A: 编辑对原稿的校对	—	3
B: 准备样张	—	2
C: 教科书封面设计	—	4
D: 准备插图	—	3
E: 作者对编辑原稿和样张的建议	A, B	2
F: 教科书排版	E	4
G: 作者对排版清样的审阅	F	2
H: 作者对插图的审阅	D	1
I: 印刷版面的生产	G, H	2
J: 教科书的印刷和装订	C, I	4

图 6.41 给出了用网络描述这些活动之间的先后关系. 虚拟活动 (2,3) 保证了同时执行的活动 A 和 B 有唯一确定的端点. 按照项目的执行进程, 用升序的数字表

示各个节点会非常方便.

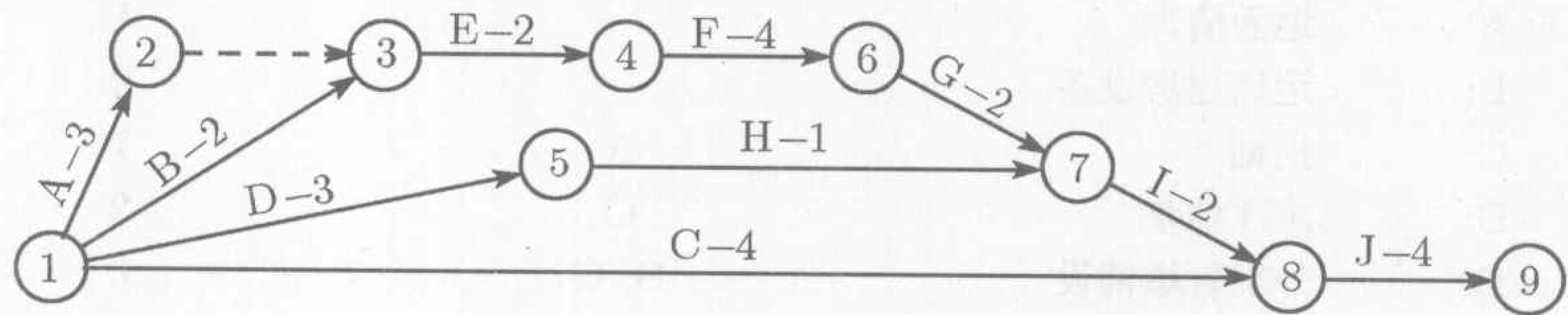


图 6.41　例 6.5-1 中的项目网络

习题 6.5 A

1. 根据下面给出的活动 A 到活动 L 之间的先后关系建立项目网络:

(a) 项目开始的 3 项活动 A, B, C 可以同时执行.

(b) A 和 B 在 D 之前.

(c) B 在 E, F, H 之前.

(d) F 和 C 在 G 之前.

(e) E 和 H 在 I 和 J 之前.

(f) C, D, F, J 在 K 之前.

(g) K 在 L 之前.

(h) I, G, L 是项目最后的 3 项活动.

2. 根据下面给出的活动 A 到 P 之间的先后关系建立项目网络:

(a) 项目开始的 3 项活动 A, B, C 可以同时执行.

(b) D, E, F 在 A 之后.

(c) I 和 G 在 B 和 D 之后.

(d) H 在 C 和 G 之后.

(e) K 和 L 在 I 之后.

(f) J 在 E 和 H 之后.

(g) M 和 N 在 F 之后, 同时必须在 E 和 H 完成之后才能执行.

(h) O 在 M 和 I 之后.

(i) P 在 J, L, O 之后.

(j) K, N, P 是项目的最后 3 项活动.

***3.** 一幢建筑的地基可以通过 4 个相继的阶段完成. 每个阶段的活动包括 (1) 挖掘; (2) 放置钢筋; (3) 灌注混凝土. 每个阶段的挖掘只有当它前一个阶段全部完成之后才能进行, 同样对灌注混凝土也是一样. 根据这个问题建立项目网络.

4. 在第 3 题中, 假定铺设管道工作的 10% 可以与第一阶段的挖掘同时进行, 但必须在灌注混凝土之前. 在地基的每一个阶段完成之后, 如果先前 5% 的铺设管道工作已经完成, 那么另外的 5% 也可以进行. 剩余的铺设管道的工作可以在项目的最后进行. 根据这个问题建立项目网络.

5. 做一项民意调查需要有: 设计和印刷调查问卷, 招募和培训人员, 选择参与者, 邮寄调查问卷, 以及分析反馈的数据. 说明这些活动的先后关系假设, 并根据这些假设建立项目网络.

6. 下表中描述了建造一个新房的各项活动. 根据这些信息建立项目网络.

活动		之前需要完成的活动	持续时间 (天数)
A:	地面清理	—	1
B:	运送建房设备	—	2
C:	挖掘	A	1
D:	灌注地基	C	2
E:	外部管道铺设	B, C	6
F:	建造房子框架	D	10
G:	安装电线	F	3
H:	铺设地板	G	1
I:	布置屋顶	F	1
J:	铺设屋内管道	E, H	5
K:	屋顶铺瓦	I	2
L:	铺设绝缘盖板	F, J	1
M:	安装门窗	F	2
N:	铺砖	L, M	4
O:	隔离墙壁和天花板	G, J	2
P:	遮盖墙壁和天花板	O	2
Q:	隔离屋顶	I, P	1
R:	内部完工	P	7
S:	外部完工	I, N	7
T:	美化	S	3

7. 某公司正在筹备投资生产一种新产品的预算, 下表给出了相关的活动以及所消耗的时间. 根据这个问题建立项目网络.

活动		之前需要完成的活动	持续时间 (天数)
A:	预计销售量	—	10
B:	研究竞争的市场	—	7
C:	设计样品和设备	A	5
D:	准备生产计划	C	3
E:	估计生产产品的成本	D	2
F:	产品定价	B, E	1
G:	准备预算	E, F	14

8. 下表列出了为一只烛光合唱团提供服务的各项活动, 根据这个问题建立项目网络.

活动		之前需要完成的活动	持续时间 (天数)
A:	选择音乐	—	2
B:	学习音乐	A	14
C:	制作拷贝和购买书籍	A	14
D:	预演	B, C	3
E:	排练	D	70
F:	租用烛光台	D	14
G:	装饰烛光台	F	1
H:	布置装饰品	D	1
I:	定购合唱团的服装	D	7

(续)

	活动	之前需要完成的活动	持续时间 (天数)
J:	检查扩音系统	D	7
K:	选择音乐声带	J	14
L:	安装扩音系统	K	1
M:	最后的排练	E, G, L	1
N:	合唱晚会	H, L, M	1
O:	最终计划	I, N	1

9. 由于道路拓宽, 需要重新铺设 1 700 英尺长的 13.8 kV 的架空基础线路. 下表列出了这项工程的一些主要活动. 根据这个问题建立项目网络.

	活动	之前需要完成的活动	持续时间 (天数)
A:	工作调查	—	1
B:	建议居民做好暂时停电准备	A	$\frac{1}{2}$
C:	物料储备	A	1
D:	审核工作	A	$\frac{1}{2}$
E:	准备电杆和材料	C, D	3
F:	分布电杆	E	$3\frac{1}{2}$
G:	电杆位置调整	D	$\frac{1}{2}$
H:	重新选择位置	G	$\frac{1}{2}$
I:	挖洞	H	3
J:	安装电杆	F, I	4
K:	覆盖旧的导线	F, I	1
L:	安装新的导线	J, K	2
M:	安装剩余设施	L	2
N:	垂下导线	L	2
O:	修整树木	D	2
P:	断开电源	B, M, N, O	$\frac{1}{10}$
Q:	接通新线电源	P	$\frac{1}{2}$
R:	整理	Q	1
S:	移走旧的导线	Q	1
T:	移走旧的电杆	S	2
U:	将剩余材料运回仓库	R, T	2

10. 下表列出了购买一辆新车时需要做的一些活动. 建立这个问题的项目网络.

	活动	之前需要完成的活动	持续时间 (天数)
A:	购买汽车之前的可行性分析	—	3
B:	找出当前汽车可能的买主	A	14
C:	列出可供购买的车型	A	1
D:	考察每一款可供购买的车型	C	3
E:	与供应商接触	C	1
F:	收集商人宣传材料	C	2
G:	汇编相关数据	D, E, F	1
H:	选择 3 款比较钟爱的车型	G	1
I:	测试这 3 款车型	H	3

(续)

	活动	之前需要完成的活动	持续时间 (天数)
J:	收集这 3 种车型的认证与价格信息	H	2
K:	选择其中一种车型	I, J	2
L:	选择供应商	K	2
M:	选择喜欢的颜色与其他参数	L	4
N:	再一次测试选中的车型	L	1
O:	购买新车	B, M, N	3

6.5.2 关键路径 (CPM) 的计算

CPM 的最终结果是要建立某个项目的活动时间表 (如图 6.38). 为了方便地达到这个目标, 我们执行一些特殊的计算来得到下面的信息:

(1) 完成一个项目总的持续时间;

(2) 将项目中的所有活动分成*关键的*和*非关键的*.

如果一项活动的开始时间和结束时间不允许有误差, 那么这个活动就称为**关键的** (critical). 而**非关键** (noncritical) 的活动则允许有计划安排方面的松弛, 以便于在不影响整个项目完成时间的前提下, 这项活动的开始时间在限制时间内可以提前或者延后.

为了执行计算中的需要, 我们定义**事件** (event) 为一项活动结束而另一项活动开始的时刻. 在网络表示中, 一个事件对应一个节点. 定义

$$\square_j = \text{事件 } j \text{ 最早的发生时间}$$

$$\triangle_j = \text{事件 } j \text{ 最晚的发生时间}$$

$$D_{ij} = \text{活动 } (i,j) \text{ 的持续时间}$$

定义指明了事件 j 相对于整个项目的*最早*和*最晚*的发生时间. 关键路径的计算过程分为两个步骤: **正推法**(forward pass) 确定事件的*最早*发生时间, **逆推法** (backward pass) 计算事件的*最晚*发生时间.

正推法(最早发生时间, $\square$) 计算过程从节点 1 开始, 递归到终止节点 n.

初始 令 $\square_1 = 0$, 表示整个项目从时刻 0 开始.

一般的第 j 步 给定节点 $p,q,\cdots,v$ 是通过入弧活动 $(p,j),(q,j),\cdots,(v,j)$ 直接连接到节点 j 上, 并且事件 (节点) $p,q,\cdots,v$ 的最早发生时间已经求出, 那么事件 j 的最早发生时间可以通过下面的式子求出:

$$\square_j = \max\{\square_p + D_{pj}, \square_q + D_{qj}, \cdots, \square_v + D_{vj}\}$$

当求出节点 n 的 $\square_n$, 正推法结束. 根据定义, $\square_j$ 代表了到节点 j 最长的路径 (持续时间).

逆推法(最晚发生时间, △) 紧接着正推法的完成, 逆推法的计算过程是从节点 n 开始到节点 1 为止.

初始 令 $\triangle_n = \square_n$, 表示项目中最后一个节点的最早发生时间与最晚发生时间相同.

一般的第 j 步 给定节点 $p, q, \cdots, v$ 是通过出弧活动 $(j,p), (j,q), \cdots, (j,v)$ 直接连接到节点 j 上, 并且事件 (节点)$p, q, \cdots, v$ 的最晚发生时间已经求出, 那么事件 j 的最晚发生时间可以通过下面的式子求出:

$$\triangle_j = \min\{\triangle_p - D_{jp}, \triangle_q - D_{jq}, \cdots, \triangle_v - D_{jv}\}$$

当求出节点 1 的 $\triangle_1$, 逆推法结束. 此时, $\triangle_1 = \square_1$.

根据上面的计算结果, 如果满足下面的条件, 那么活动 (i,j) 就是**关键的**.

(1) $\triangle_i = \square_i$

(2) $\triangle_j = \square_j$

(3) $\triangle_j - \triangle_i = \square_j - \square_i = D_{ij}$

这 3 个条件说明了端点 i 和端点 j 的最早和最晚发生时间相等, 并且持续时间 D_{ij} 严格地满足指定的时间限制. 所有不满足这 3 个条件的活动都称为**非关键的**.

根据定义可以知道, 网络中关键的活动一定可以组成一条不间断的路径, 从开始到终止贯穿整个网络.

例 6.5-2

确定图 6.42 所示的项目网络中的关键路径. 所有持续时间按天数计算.

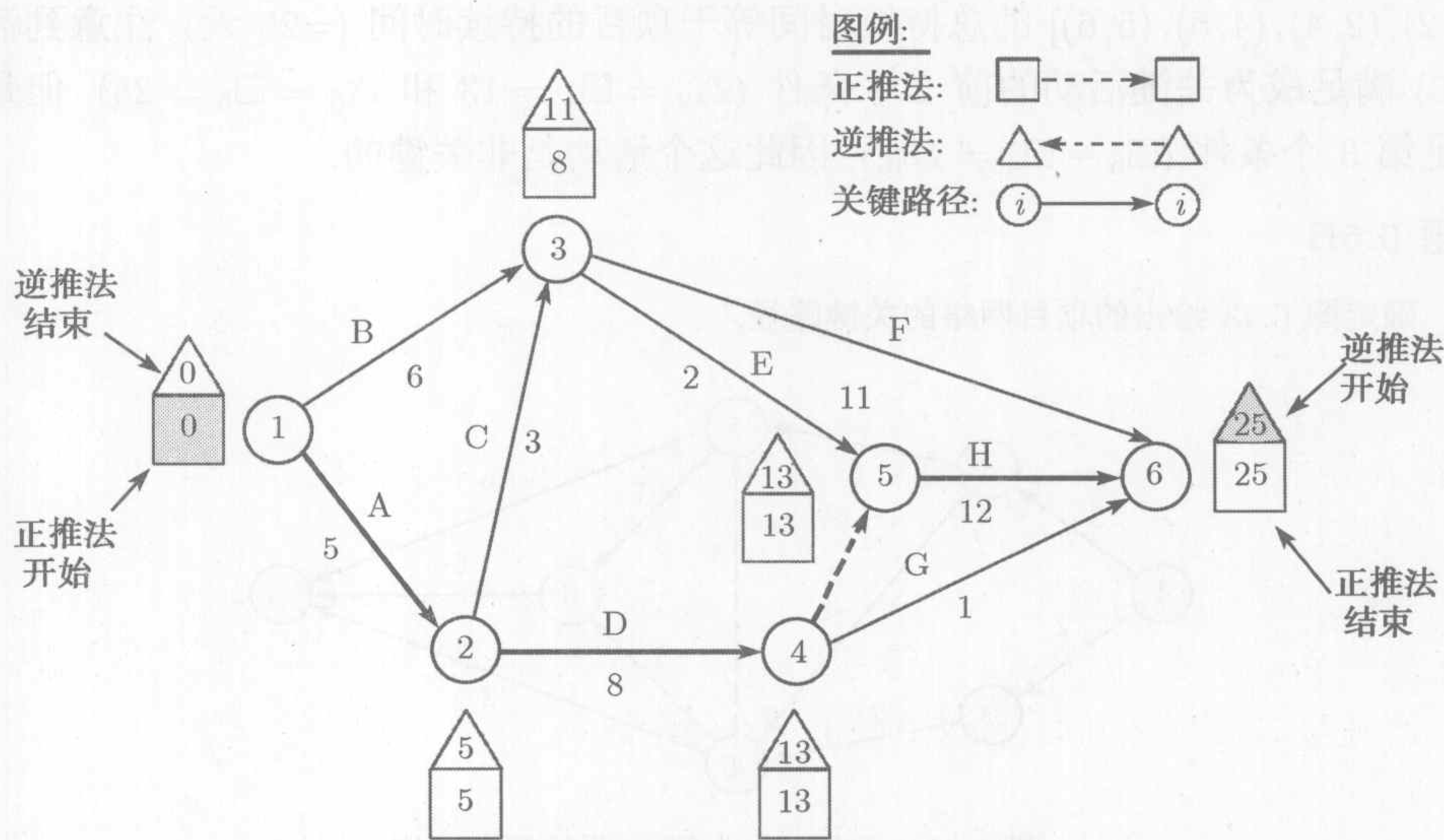


图 6.42 例 6.5-2 中项目的正推法和逆推法的计算过程

正推法

节点 1　令 $\square_1 = 0$

节点 2　$\square_2 = \square_1 + D_{12} = 0 + 5 = 5$

节点 3　$\square_3 = \max\{\square_1 + D_{13}, \square_2 + D_{23}\} = \max\{0 + 6, 5 + 3\} = 8$

节点 4　$\square_4 = \square_2 + D_{24} = 5 + 8 = 13$

节点 5　$\square_5 = \max\{\square_3 + D_{35}, \square_4 + D_{45}\} = \max\{8 + 2, 13 + 0\} = 13$

节点 6　$\square_6 = \max\{\square_3 + D_{36}, \square_4 + D_{46}, \square_5 + D_{56}\}$

$= \max\{8 + 11, 13 + 1, 13 + 12\} = 25$

计算结果显示项目可以在 25 天内完成.

逆推法

节点 6　令 $\triangle_6 = \square_6 = 25$

节点 5　$\triangle_5 = \triangle_6 - D_{56} = 25 - 12 = 13$

节点 4　$\triangle_4 = \min\{\triangle_6 - D_{46}, \triangle_5 - D_{45}\} = \min\{25 - 1, 13 - 0\} = 13$

节点 3　$\triangle_3 = \min\{\triangle_6 - D_{36}, \triangle_5 - D_{35}\} = \min\{25 - 11, 13 - 2\} = 11$

节点 2　$\triangle_2 = \min\{\triangle_4 - D_{24}, \triangle_3 - D_{23}\} = \min\{13 - 8, 11 - 3\} = 5$

节点 1　$\triangle_1 = \min\{\triangle_3 - D_{13}, \triangle_2 - D_{12}\} = \min\{11 - 6, 5 - 5\} = 0$

最终计算结果总是 $\triangle_1 = 0$.

正推法和逆推法的整个计算过程都可以在网络上直接进行, 如图 6.42 所示. 根据确定关键路径的规则, 可以得到网络的关键路径是 $1 \to 2 \to 4 \to 5 \to 6$, 这正是所需要的贯穿了从开始 (节点 1) 到终止 (节点 6) 的整个网络. 关键活动 $[(1,2),(2,4),(4,5),(5,6)]$ 的总持续时间等于项目的持续时间 (=25 天). 注意到活动 $(4,6)$ 满足成为关键活动的前 2 个条件 ($\triangle_4 = \square_4 = 13$ 和 $\triangle_5 = \square_5 = 25$), 但是不满足第 3 个条件 ($\square_6 - \square_4 \neq D_{46}$), 因此这个活动是非关键的.

习题 6.5B

*1. 确定图 6.43 给出的项目网络的关键路径.

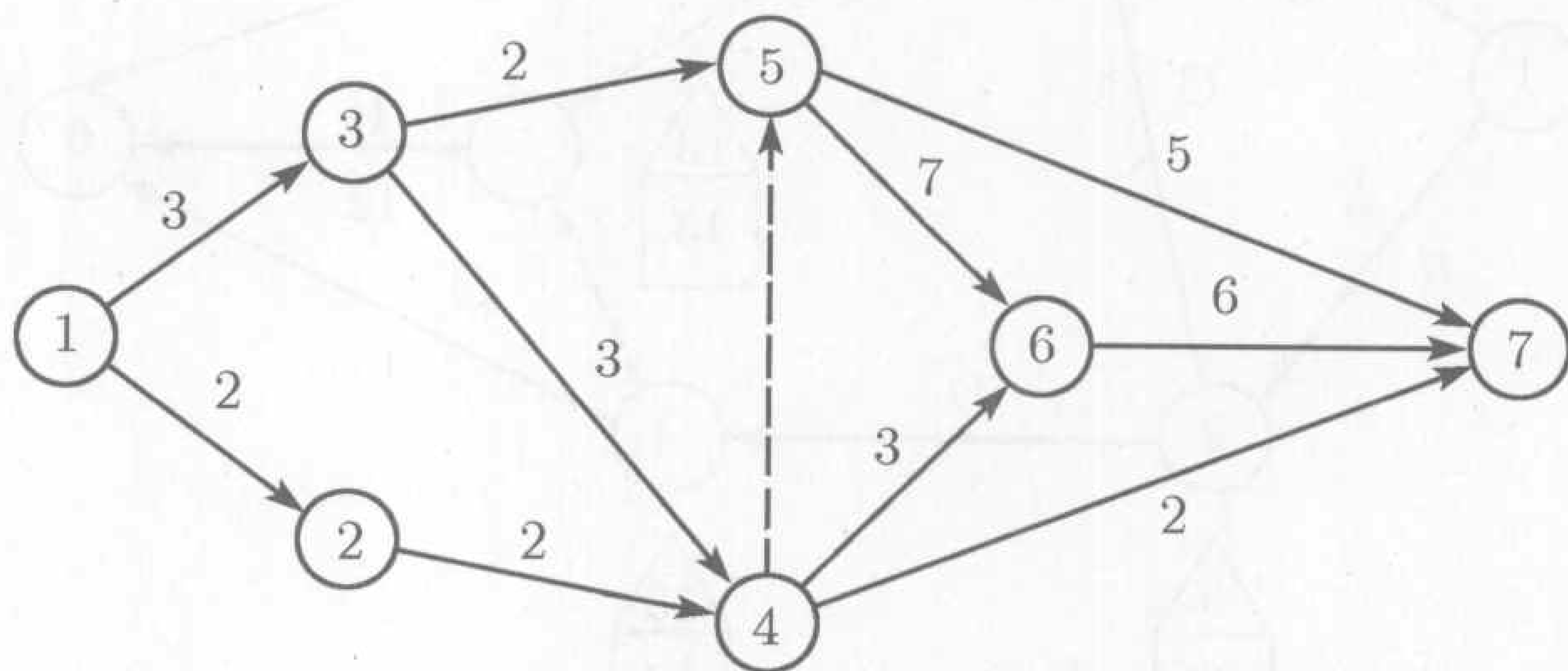


图 6.43　习题 6.5B 第 1 题的项目网络

2. 确定图 6.44 给出的项目网络的关键路径.

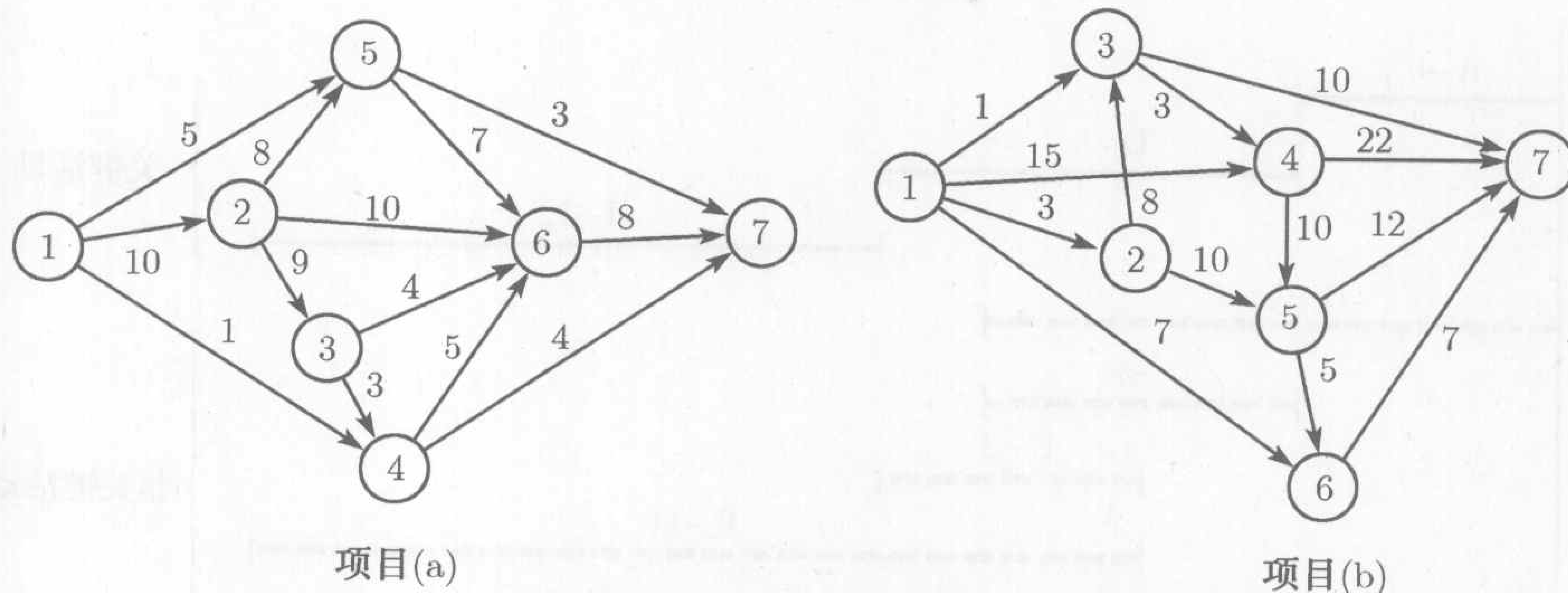


图 6.44 习题 6.5B 第 2 题的项目网络

3. 确定习题 6.5A 第 6 题中项目的关键路径.

4. 确定习题 6.5A 第 8 题中项目的关键路径.

5. 确定习题 6.5A 第 9 题中项目的关键路径.

6. 确定习题 6.5A 第 10 题中项目的关键路径.

6.5.3 建立时间表

这一节中介绍如何根据 6.5.2 节中得到的计算结果建立活动的时间表. 对活动 (i,j), $\square_i$ 表示*最早开始时间*, $\triangle_j$ 表示*最晚完成时间*. 那么区间 $(\square_i, \triangle_j)$ 代表了在不延误整个项目的前提下, 活动 (i,j) 可以安排进行的最大跨度区间.

建立初步的时间表 我们通过一个例题来讲解如何建立一个初步的时间表.

例 6.5-3

确定例 6.5-2(图 6.42) 中项目的时间表.

我们可以得到项目中不同活动的初步时间表, 用来描绘各个活动各自的时间跨度区间, 如图 6.45 所示, 从图中可以得到下面的两点结论.

(1) 关键活动 (图中实线表示) 一定是从左到右一个接着一个地排列, 使得项目的完成时间在指定的 25 天内.

(2) 非关键活动 (图中虚线表示) 有比它们各自持续时间要长的一个时间跨度区间, 因此在时间表中在它们各自的时间跨度区间内允许有时间延迟 (或者误差).

对于每一个非关键活动如何在它允许的时间跨度区间上建立时间表? 一般地, 人们希望每一个非关键活动都可以尽可能早地进行. 按照这种方法, 松弛阶段还是会有助于在安排完活动之后来减少执行活动中不必要的延误时间. 然而, 也会有某个非关键的活动只能在它的最早允许开始时间之后执行的情况. 例如, 在图 6.45 中, 假定每一个非关键活动 E 和 F 执行时都需要 1 台推土机, 然而当前只有 1 台推土机, 所以 E 和 F 不能同时进行, 要想同时在最早的第 8 天到第 10 天进行则需要 2 台推土机. 因此为了避免冲突, 只能首先将活动 E 的开始时间定在第 8 天, 然后把活动 F 的开始时间放在第 10 天到第 14 天之间.

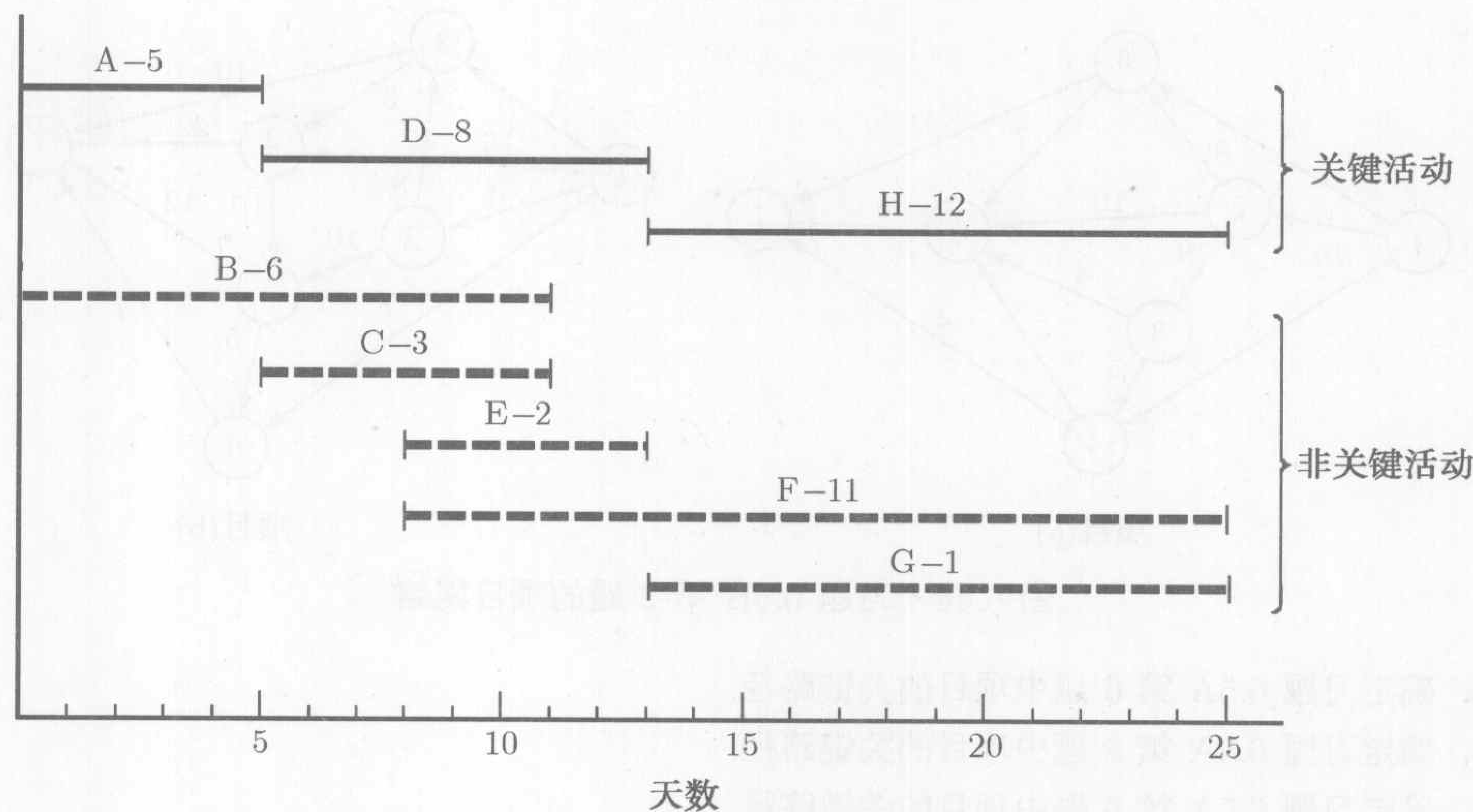


图 6.45　例 6.5-2 中项目的初步时间表

如果所有的非关键活动都可以按照各自最早的开始时间进行安排, 那么得到的时间表当然是可行的. 否则, 如果某个非关键活动的开始时间不是它的最早开始时间, 那么就可能使得活动之间的关系受到破坏. 例如, 图 6.45 中的活动 C 和 E. 在项目网络中 (图 6.42), 虽然 C 必须在 E 之前完成, 但是从图 6.45 中 C 和 E 的时间跨度区间可以看出, 活动 C 可以安排在第 6 天到第 9 天之间, E 可以安排在第 8 天到第 10 天之间, 这会违背 C 在 E 之前的要求. 需要用一面“红旗”来自动地标示出时间表的冲突, 这样的信息需要计算每一项非关键活动的*浮动数*.

浮动数的确定　浮动数表示非关键活动的所分配跨度区间内的可用松弛时间, 最普通的是**总浮动数** (total float) 和**自由浮动数** (free float).

图 6.46 给出了计算一项活动 (i,j) 的总浮动数 (TF_{ij}) 和自由浮动数 (FF_{ij}) 的简便方法. 活动 (i,j) 的总浮动数就是, 从事件 i *最早发生*时间到事件 j *最晚发生*时间的跨度区间减去活动 (i,j) 的持续时间, 也就是,

$$TF_{ij} = \triangle_j - \square_i - D_{ij}$$

自由浮动数是从事件 i *最早发生*时间到事件 j *最早发生*时间的跨度区间减去活动 (i,j) 的持续时间, 也就是,

$$FF_{ij} = \square_j - \square_i - D_{ij}$$

根据定义可知 $FF_{ij} \leqslant TF_{ij}$.

标记红旗的原则　*对于一个非关键活动* (i,j),

(a) *当* $FF_{ij} = TF_{ij}$ *时, 可以将这项活动安排在它的跨度区间* $(\square_j, \triangle_j)$ *中的任何位置, 而不会造成时间冲突.*

(b) 当 $FF_{ij} < TF_{ij}$ 时, 这项活动的开始时刻相对于它的最早开始时间 ($\square_i$) 而言, 在不会造成排序冲突的前提下最多可以延迟 FF_{ij}. 如果延迟时间超过了 FF_{ij}(同时不超过 TF_{ij}), 那么离开节点 j 的所有活动的开始时间一定也会延迟同样长的时间.

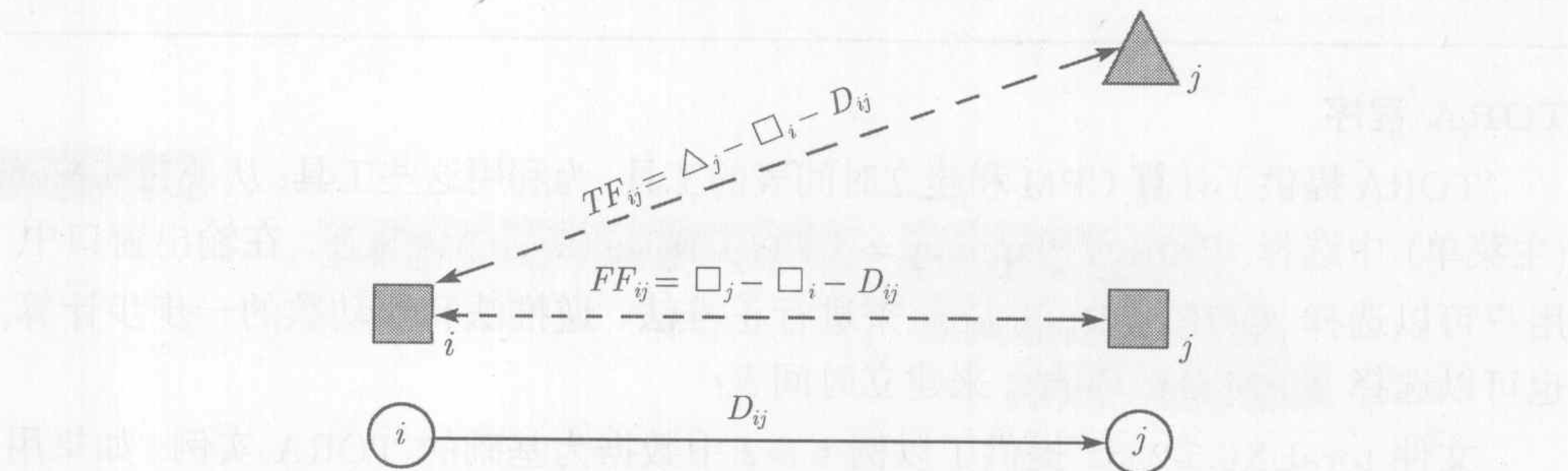


图 6.46 总浮动数和自由浮动数的计算

这个原则就暗示出如果一个非关键活动 (i, j) 满足 $FF_{ij} < TF_{ij}$, 那么它将被标记红旗. 红旗标记的重要性体现在, 对于需要延误某个活动的开始时间超出它的最早开始时间 ($\square_i$) 的情况, 在这种情况下, 为了避免时间冲突, 需要特别注意离开节点 j 的所有其他活动的开始时间.

例 6.5-4

计算例 6.5-2 网络中非关键活动的浮动数, 并讨论它们在最终得到项目活动时间表中的用途.

下表汇总了总浮动数和自由浮动数的计算结果. 也可以更加方便地直接在图 6.42 中的网络上计算.

非关键活动	持续时间	总浮动数(TF)	自由浮动数 (FF)
B(1,3)	6	$11-0-6=5$	$8-0-6=2$
C(2,3)	3	$11-5-3=3$	$8-5-3=0$
E(3,5)	2	$13-8-2=3$	$13-8-2=3$
F(3,6)	11	$25-8-11=6$	$25-8-11=6$
G(4,6)	1	$25-13-1=11$	$25-13-1=11$

根据计算结果由于活动 B 和 C 对应的 $FF < TF$, 所有它们被标记红旗. 剩余的活动 (E, F, G) 满足 $FF = TF$, 所以它们可以分配在相应最早开始时间和最晚完成时间中的任何一个时刻开始.

为了演示标记红旗活动的重要性, 我们以活动 B 为例. 因为它的 $TF = 5$ 天, 那么这个活动的开始时间最早可以在时刻 0, 最晚在第 5 天 (如图 6.45 所示). 同时又因为 $FF = 2$ 天, 所以活动 B 的开始时间可以在 0 到 2 的任何一个时刻, 同时不会影响接下来的活动 E 和 F. 然而, 如果活动 B 必须在时刻 $2 + d(\leqslant 5)$ 开始, 那么

紧接着的两个活动 E 和 F 的开始时间一定要比它的最早开始时间 (= 8) 至少推迟 d. 按照这种方法, 就可以保持活动 B 与活动 E 和 F 的相继关系.

再看红旗标记的活动 C, 注意到它对应的 $FF = 0$. 这就意味着活动 C 相对于它最早开始时间的任何长度的延迟, 都会导致活动 E 和 F 至少同样长度的延迟.

TORA 程序

TORA 提供了计算 CPM 和建立时间表的工具, 为利用这些工具: 从 Main Menu (主菜单) 中选择 Project Planning ⇒ CPM-Critical Path Method. 在输出窗口中, 用户可以选择 CMP Calculations 来进行正推法、逆推法和浮动数的一步步计算, 也可以选择 CPM Bar Chart 来建立时间表.

文件 toraEX6.5-2.txt 提供了以例 6.5-2 中数据为基础的 TORA 实例. 如果用户选择应用 Next Step 选项来产生输出, TORA 将会提供详细的计算正推法和逆推法的过程.

图 6.47 给出了对于例 6.5-2 中项目的数据, 选择 CPM Bar Chart 产生的时间表. 默认生成的图表是对所有的非关键活动做出的尽量早的安排. 当然读者也可以利用屏幕左下方的自解释列表, 来更多地了解如果延迟一个非关键活动的开始时间, 将会有什么样的影响. 这种延迟影响可以通过带有解释语句的图表直接显示出来. 例如, 如果活动 B 的开始时间延误了超过 2 天, 紧接着的活动 E 和 F 的延迟时间就等于 B 延误时间与它自由浮动数之间的差. 特别地, 已知 B 的自由浮动数是 2 天, 如果 B 的延迟时间是 3 天, 那么 E 和 F 的延迟时间至少是 $3 - 2 = 1$ 天, 如图 6.47 给出了详细的结果.

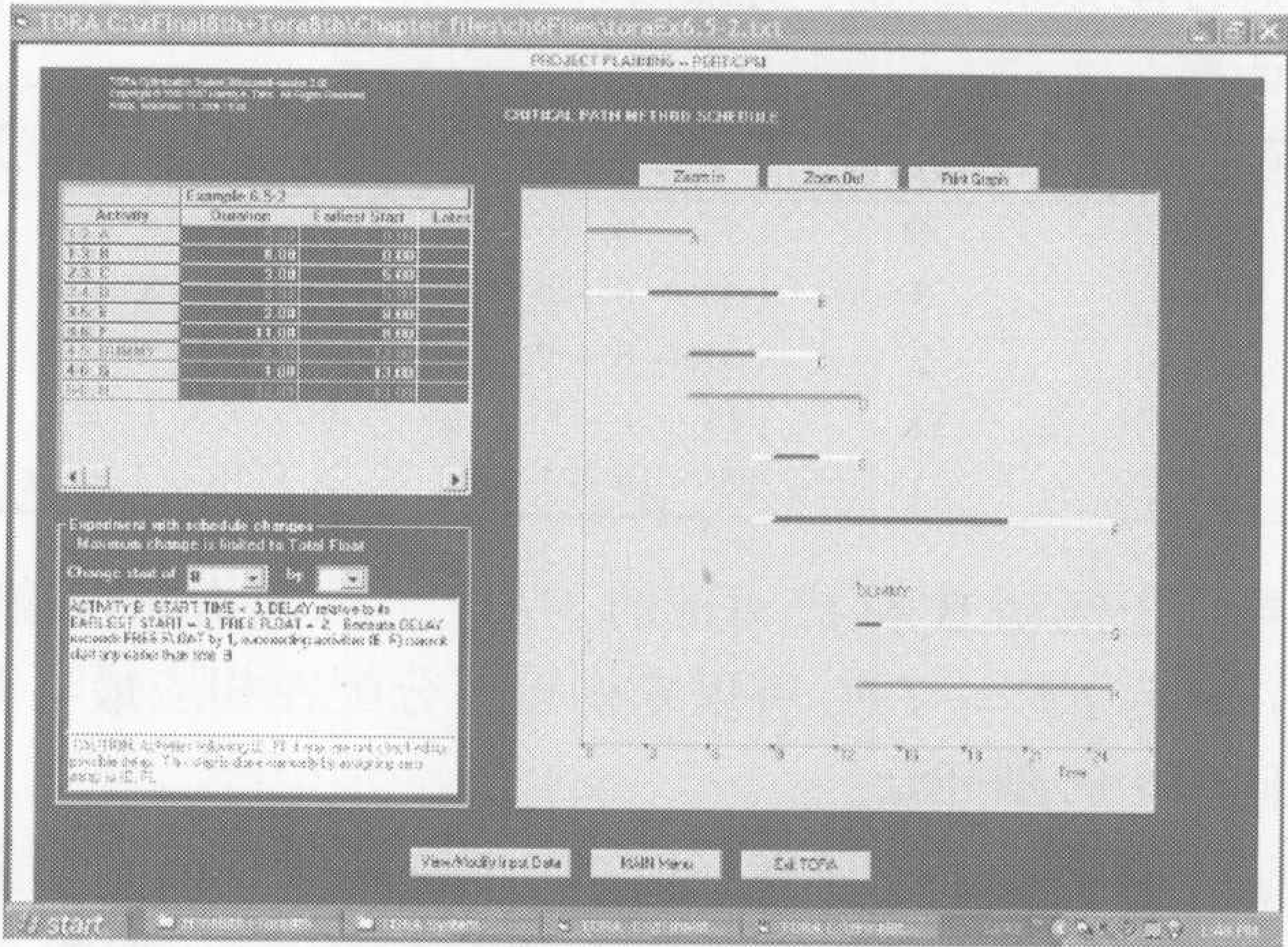


图 6.47　TORA 求解例 6.5-2 得到的结果 (文件 toraEx6.5-2.txt)

```
#---------- CPM (Example 6.5-2) ----------
paramn;
param D{1..n,1..n} default -1;

set into{1..n};
set from{1..n};

var x{i in 1..n,j in from[i]}>=0;
var ET{i in 1..n};
var LT{i in 1..n};
var TF{i in 1..n, j in from[i]};
var FF{i in 1..n, j in from[i]};

data;
param n:=6;
param D: 1 2 3 4 5 6:=
1 . 5 6 . . .
2 . . 3 8 . .
3 . . . . 2 11
4 . . . . 0 1
5 . . . . . 12
6 . . . . . .;

for {i in 1..n} {let from[i]:={j in 1..n:D[i,j]>=0}};
for {j in 1..n} {let into[j]:={i in 1..n:D[i,j]>=0}};

#----- nodes earliest and latest times and floats
let ET[1]:=0;          #earliest node time
for {i in 2..n}let ET[i]:=max{j in into[i]}(ET[j]+D[j,i]);

let LT[n]:=ET[n];      #latest node time
for{i in n-1..1 by -1}let LT[i]:=min{j in from[i]}(LT[j]-D[i,j]);

printf "%1s-%1s %5s %5s %5s %5s %5s %5s %5s \n\n",
"i","j","D","ES","EC","LS","LC","TF","FF" >Ex6.6-2out.txt;
for {i in 1..n, j in from[i]}
{
let TF[i,j]:=LT[j]-ET[i]-D[i,j];
let FF[i,j]:=ET[j]-ET[i]-D[i,j];
printf "%1i-%1i %5i %5i %5i %5i %5i %5i %5i %3s\n",
i,j,D[i,j],ET[i],ET[i]+D[i,j],LT[j]-D[i,j],LT[j],TF[i,j],FF[i,j],
if TF[i,j]=0 then "c" else " " >Ex6.6-2out.txt;
}
```

图 6.48 求解例 6.5-2 的 AMPL 模型 (文件 amplEx6.5-2.txt)

AMPL 程序

图 6.48 提供了求解 CPM 的 AMPL 模型 (文件 amplEx6.5-2.txt), 这个模型使用的是例 6.5-2 的数据. 因为这个 AMPL 模型使用了索引集合并且不需要优化, 所以它是唯一的. 实质上, 不需要 `solver` 命令, AMPL 的实现非常类似与纯的程序语言 Basic 或者 C.

在计算 CPM 的时候, 为了表示网络需要对每一个节点引入两个索引集: `into`和`from`. 对于节点`i`, 集`into[i]`表示所有流向`i`的节点集合, 集`from[i]`表示所有从`i`可以到达的节点集合. 例如, 在例 6.5-2 中, `from[1]={2,3}`, `into[1]`是空集.

可以通过下面的方法得到模型中的子集`from`和`into`: 由于 CPM 网络引入虚拟活动时`D[i,j]`可以是 0, 因此对于网络中所有不存在的弧设定`D[i,j]`的默认值为 −1. 集`from[i]`就表示集`{1..n}`中所有满足`D[i,j]>=0`的节点`j`组成的集合. 也就是说集`from[i]`可以定义为子集`{j in 1..n:D[i,j]>=0}`. 同样的道理可以用于确定子集`into[i]`. 给定了数据`D[i,j]`, 如图 6.48 所示, AMPL 模型利用下面两条语句, 自动地确定出这两个集:

```
for {i in 1..n} {let from[i]:={j in 1..n:D[i,j]>=0}};
for {j in 1..n} {let into[j]:={i in 1..n:D[i,j]>=0}};
```

一旦确定出了集`from`和`into`, 模型利用正推法可以计算出最早开始时间, `ET[i]`. 在这之后, 可以通过下面的语句初始化逆推法:

```
let LT[n]:=ET[n];
```

模型剩余的任务就是要得到如图 6.49 所示的输出结果. 这个输出结果给出了绘制 CPM 图表所需要的所有数据, 这一部分利用的是例 6.5-2 和例 6.5-4 的计算结果.

```
i-j    D   ES   EC   LS   LC   TF   FF

1-2    5    0    5    0    5    0    0   c
1-3    6    0    6    5   11    5    2
2-3    3    5    8    8   11    3    0
2-4    8    5   13    5   13    0    0   c
3-5    2    8   10   11   13    3    3
3-6   11    8   19   14   25    6    6
4-5    0   13   13   13   13    0    0   c
4-6    1   13   14   24   25   11   11
5-6   12   13   25   13   25    0    0   c
```

图 6.49 例 6.5-2 对应的 AMPL 模型的输出结果 (文件 amplEx6.5-2.txt)

习题 6.5C

1. 给定一项持续时间为 D_{ij} 的活动 (i, j), 以及它的最早开始时间 $\square_i$ 和最晚完成时间 $\triangle_j$, 求这项活动的最早完成时间和最晚开始时间.
2. 一项关键活动的总浮动数和自由浮动数的含义是什么? 并给出解释.

***3.** 对于下面的每一项活动, 在允许紧接着它的活动可以安排在相应最早和最晚完成时间内任意一个时刻的情况下, 确定这项活动开始时间与它最早开始时间的最大延迟.

(a) $TF = 10, FF = 10, D = 4$　　**(b)** $TF = 10, FF = 5, D = 4$

(c) $TF = 10, FF = 0, D = 4$

4. 在例 6.5-4 中, 利用浮动数回答下面的问题:

(a) 如果活动 B 在第 1 天开始, 活动 C 在第 5 天开始, 确定活动 E 和 F 的最早开始时间.

(b) 如果活动 B 在第 3 天开始, 活动 C 在第 7 天开始, 确定活动 E 和 F 的最早开始时间.

(c) 如果活动 B 在第 6 天开始, 这会对其他活动有什么样的影响?

***5.** 在例 6.5-2 的项目中 (如图 6.42 所示), 假定活动 B 和活动 F 的持续时间分别从 6 天增加到 11 天和从 20 天增加到 25 天.

(a) 确定关键路径.

(b) 确定网络中的总浮动数和自由浮动数, 并指出标记红旗的活动.

(c) 如果活动 A 的开始时间为 5, 确定活动 C, D, E, G 的最早可能开始时间.

(d) 如果活动 F, G, H 都需要同一种设备, 确定需要这种设备的最小数目.

6. 计算图 6.44 中项目 (a) 和 (b) 的浮动数并指出标记红旗的活动. 然后针对下面的情况确定时间表:

项目(a)

(i) 活动 $(1,5)$ 不能在第 14 天之前开始.

(ii) 活动 $(5,6)$ 和 $(5,7)$ 需要使用同一种设备, 并且这种设备只有一套可用.

(iii) 所有其他活动都尽量早地开始.

项目(b)

(i) 活动 $(1,3)$ 必须安排在它的最早开始时间, 同时活动 $(1, 2)$、$(1, 3)$ 和 $(1, 6)$ 使用同一种特殊的设备, 并且这种设备只有一套可用.

(ii) 所有其他活动都尽量早地开始.

6.5.4 CPM 的线性规划模型

CPM 问题可以看作是最短路径问题 (6.3 节) 的反问题, 实质上就是找到一条最长的路径, 将一个单位的流从开始点流入并从终点流出, 所以可以利用 6.3.3 节中建立最短路径问题线性规划模型的方法, 定义

$$x_{ij} = \text{活动 } (i,j) \text{ 上的流量, 对所有定义的 } i \text{ 和 } j$$
$$D_{ij} = \text{活动 } (i,j) \text{ 的持续时间, 对所有定义的 } i \text{ 和 } j$$

那么, 线性规划的目标函数是

$$\max\ z = \sum_{\text{所有已定义的活动}(i,j)} D_{ij}x_{ij}$$

(与最短路径问题的线性规划模型的不同之处就是目标函数不再是最小化.) 对于每一个节点, 需要一个约束来保证流的守恒:

$$总的流入量 = 总的流出量$$

所有的变量 x_{ij} 都是非负的.

例 6.5-5

下面给出了例 6.5-2 项目问题 (图 6.42) 的线性规划模型, 其中节点 1 和节点 6 分别是开始点和终止点.

	A	B	C	D	E	F	虚拟活动	G	H	
	x_{12}	x_{13}	x_{23}	x_{24}	x_{35}	x_{36}	x_{45}	x_{46}	x_{56}	
max $z=$	6	6	3	8	2	11	0	1	12	
节点 1	−1	−1								=−1
节点 2	1		−1	−1						=0
节点 3		1	1		−1	−1				=0
节点 4				1			−1	−1		=0
节点 5					1		1		−1	=0
节点 6						1		1	1	=1

最优解是

$$z = 25, x_{12}(\text{A}) = 1, x_{24}(\text{D}) = 1, x_{45}(\text{虚拟活动}) = 1, x_{56}(\text{H}) = 1, \text{其他的} = 0$$

根据这个解得到的关键路径是 A → D → 虚拟活动 → H, 并且项目的持续时间是 25 天. 线性规划模型的解并不完整, 因为它只给出了关键路径, 而没有提供建立 CPM 图表的数据, 然而我们已经从图 6.48 知道, AMPL 可以在不求解线性规划最优的情况下提供所有需要的数据.

习题 6.5D

1. 利用线性规划模型确定图 6.43 中项目网络的关键路径.
2. 利用线性规划模型确定图 6.44 中项目网络的关键路径.

6.5.5 PERT 网络

与 CPM 不同的是, PERT 基于对活动持续时间的 3 个估计:

(1) **最乐观时间** a, 在条件最好的情况下执行所需要的时间.

(2) **最可能时间** m, 在条件正常的情况下执行所需要的时间.

(3) **最悲观时间** b, 在条件最差的情况下执行所需要的时间.

区间 (a, b) 包含了所有可能的活动持续时间的估计, 估计值 m 在区间 (a, b) 上. 根据这些估计, 那么平均的持续时间 $\overline{D}$ 和方差 v 可以近似为:

$$\overline{D}=\frac{a+4m+b}{6}, \quad v=\left(\frac{b-a}{6}\right)^2$$

可以直接利用 6.5.2 节和 6.5.3 节中 CPM 的计算过程, 用 $\overline{D}$ 代替单一的估计值 D.

下面就可以估计网络中一个节点 j 在事先指定的调度时间 S_j 之前出现的概率. 令 e_j 表示节点 j 最早发生的时间. 由于从开始节点到节点 j 之间的活动的持续时间是一个随机变量, 所以 e_j 也是一个随机变量. 假定网络中所有的活动都是统计上独立的, 我们可以按照下面的方法确定均值 $\mathrm{E}\{e_j\}$ 和方差 $\mathrm{var}\{e_j\}$. 如果从开始节点到节点 j 只有一条路径, 那么均值就等于这条路径上所有活动的期望持续时间 $\overline{D}$ 之和, 方差等于这些活动的方差 v 之和. 但是, 如果从开始节点到节点 j 有多于一条的路径, 那么首先需要确定最长路径持续时间的统计分布, 这是一个非常难的问题, 因为它等同于确定两个或者多个随机变量中最大的一个变量的分布. 所以在计算均值 $\mathrm{E}\{e_j\}$ 和方差 $\mathrm{var}\{e_j\}$ 时, 我们有一个简化的假设, 那就是在所有到节点 j 的路径中选择活动期望持续时间之和最大的一条. 如果两条以上的路径拥有相同的均值, 那么就取其中方差最大的一条, 因为这样做可以更好地反映出不确定性, 也就能够得到概率的一个比较保守的估计.

一旦求出了一条到节点 j 的路径上的均值 $\mathrm{E}\{e_j\}$ 和方差 $\mathrm{var}\{e_j\}$, 节点 j 在预先设定的时间 S_j 之前出现的概率就可以按照下面的式子求出:

$$\mathrm{P}\{e_j\leqslant S_j\}=\mathrm{P}\left\{\frac{e_j-\mathrm{E}\{e_j\}}{\sqrt{\mathrm{var}\{e_j\}}}\leqslant\frac{S_j-\mathrm{E}\{e_j\}}{\sqrt{\mathrm{var}\{e_j\}}}\right\}=\mathrm{P}\{z\leqslant K_j\}$$

其中,

$$z=\text{标准正态随机变量}, \quad K_j=\frac{S_j-\mathrm{E}\{e_j\}}{\sqrt{\mathrm{var}\{e_j\}}}$$

标准正态随机变量 z 的均值是 0, 标准差是 1(见 14.4.4 节). 由于 e_j 是独立随机变量之和, 所以应用正态分布是合理的. 根据**大数定律与中心极限定理** (见 14.4.4 节), e_j 近似地服从正态分布.

例 6.5-6

考虑例 6.5-2 中的项目. 为了避免关键路径的重复计算, 按照下表选择 a, m 和 b 的值, 使得例 6.5-2 中所有的 i 和 j 满足 $\overline{D}_{ij}=D_{ij}$.

活动	$i-j$	(a,m,b)	活动	$i-j$	(a,m,b)
A	1 − 2	(3,5,7)	E	3 − 5	(1,2,3)
B	1 − 3	(4,6,8)	F	3 − 6	(9,11,13)
C	2 − 3	(1,3,5)	G	4 − 6	(1,1,1)
D	2 − 4	(5,8,11)	H	5 − 6	(10,12,14)

下表给出了不同活动的均值 $\overline{D}_{ij}$ 和方差 v_{ij}. **注意到对于一个虚拟活动有** $(a,m,b)=(0,0,0)$, 因此它的均值和方差都是 0.

活动	$i-j$	$\overline{D}_{ij}$	v_{ij}	活动	$i-j$	$\overline{D}_{ij}$	v_{ij}
A	1 − 2	5	0.444	E	3 − 5	2	0.111
B	1 − 3	6	0.444	F	3 − 6	11	0.444
C	2 − 3	3	0.444	G	4 − 6	1	0.000
D	2 − 4	8	1.000	H	5 − 6	12	0.444

下面一个表格给出了从节点 1 到其他不同节点的最长路径, 以及相应的均值和标准差.

节点	基于持续时间均值的最长路径	路径均值	路径标准差
2	1 2	5.00	0.67
3	1 − 2 − 3	8.00	0.94
4	1 − 2 − 4	13.00	1.20
5	1 − 2 − 4 − 5	13.00	1.20
6	1 − 2 − 4 − 5 − 6	25.00	1.37

最后一张表格计算的是按照分析指定的每个节点在时间 S_j 之前出现的概率.

节点 j	最长路径	路径均值	路径标准差	S_j	K_j	$P\{z \leqslant K_j\}$
2	1 − 2	5.00	0.67	5.00	0	0.500 0
3	1 − 2 − 3	8.00	0.94	11.00	3.19	0.999 3
4	1 − 2 − 4	13.00	1.20	12.00	−0.83	0.203 3
5	1 − 2 − 4 − 5	13.00	1.20	14.00	0.83	0.796 7
6	1 − 2 − 4 − 5 − 6	25.00	1.37	26.00	0.73	0.767 3

TORA 模型

TORA 提供了求解 PERT 的模板. 应用这个模板的方法是, 在 Main Menu 中选择 Project Planning ⇒ PERT-Program Evaluation and Review Technique .在输出窗口, 可以选择 Activity Mean/Var 来计算每一项活动的均值和方差, 或者选择 PERT Calculations 来计算网络中到每个节点最长路径的均值和方差. 文件 toraEx 6.5-6.txt 是求解例 6.5-6 的 TORA 例子.

习题 6.5E

1. 考虑习题 6.5B 第 2 题. 下表给出了对应 (a, m, b) 的估计. 求项目中不同节点在没有延迟的情况下执行的概率.

项目 (a)				项目 (b)			
活动	(a, m, b)	活动	(a, m, b)	活动	(a, m, b)	活动	(a, m, b)
1 − 2	(5,6,8)	3 − 6	(3,4,5)	1 − 2	(1,3,4)	3 − 7	(12,13,14)
1 − 4	(1,3,4)	4 − 6	(4,8,10)	1 − 3	(5,7,8)	4 − 5	(10,12,15)
1 − 5	(2,4,5)	4 − 7	(5,6,8)	1 − 4	(6,7,9)	4 − 7	(8,10,12)
2 − 3	(4,5,6)	5 − 6	(9,10,15)	1 − 6	(1,2,3)	5 − 6	(7,8,11)
2 − 5	(7,8,10)	5 − 7	(4,6,8)	2 − 3	(3,4,5)	5 − 7	(2,4,8)
2 − 6	(8,9,13)	6 − 7	(3,4,5)	2 − 5	(7,8,9)	6 − 7	(5,6,7)
3 − 4	(5,9,19)			3 − 4	(10,15,20)		

参考文献

Ahuja, R., T. Magnati, and J. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, Upper Saddle River, NJ, 1993.

Bazarra, M., J. Jarvis, and H. Sherali, *Linear Programming and Network Flow*, 2nd ed., Wiley, New York, 1990.

Bersetkas, D., *Network Optimization: Continuous and Discrete Models*, Athena Scientific, Nashua, NH, 1998.

Charnes, A. and W. Cooper, "Some Network Characterization for Mathematical Programming and Accounting Applications to Planning and Control," *The Accounting Review*, Vol. 42, No. 3,pp. 24-52, 1967.

Evans, J. R., and E. Minieka, *Optimization Algorithms for Networks and Graphs*, 2nd ed., Marcel Dekker, New York, 1992.

Guéret, C., C. Prins, and M. Sevaux, *Applications of Optimization with Xpress-MP*, translated and revised by Susanne Heipke, Dash Optimization Ltd., London, 2002.

Glover, F., D. Klingman, and N. Phillips, *Network Models and Their Applications in Practice*, Wiley, New York, 1992.

Glover, F., and M. Laguna, *Tabu Search*, Kulwer Academic Publishers, Boston, 1997.

Murty, K., *Network Programming*, Prentice Hall, Upper Saddle River, NJ, 1992.

Robinson, E. W, L. Gao, and S. Muggenborg, "Designing an Integrated Distribution System at DowBrands, Inc," *Interfaces*, Vol. 23, No. 3, pp. 107-117, 1993.

第7章　目标规划

本章导读　前几章所介绍的线性规划模型是基于优化单个目标函数的模型. 但是在有些情形下, 往往要求优化多个目标, 甚至其中的某些目标是相互冲突的. 例如, 政府希望在减少财政赤字的同时, 减少个人所得税, 在这样的一些情形下, 找到单个解使它同时能够优化这样两个相互冲突的目标是不可能的. 目标规划所要做的就是, 根据各个目标的相对重要程度, 试图寻找一个*折中*的解.

本章主要需要的预备知识是对单纯形方法的基本了解. 解决目标规划有两种方法: *权和法*和*设定优先权法*. *权和法*是构造一个新的单一目标函数, 它是由各个目标加权求和得到的; *设定优先权法*是从具有最高优先级的目标开始逐个优化, 直到最低优先级的目标, 并且每次优化均不会降低较高优先级的目标值. 权和法的求解方法就像一个普通的线性规划. 设定优先权法则还需要考虑在第 3 章中的单纯形算法基础上的"额外"算法. 本章还将介绍 AMPL 模型, 它通过简单地修改所输入的数据, 就可以很容易地把设定优先权法交互应用到任一目标规划的求解中. 希望读者认真学习这个模型, 因为它会有助于更好地理解设定优先权法的思想.

本章包括 1 个实际应用的综述、4 个例题、1 个 AMPL 模型、25 个节后习题和 2 个案例. 所有的案例都放在附录 E 中, AMPL/Excel/Solver/TORA 的程序放在文件夹 ch7Files 中.

实际应用——Mount Sinai 医院中的手术室调配时间表问题

这种情况出现在加拿大, 在那里医疗保险是强制性的和全民性的. 每个省自行控制主要以保险费和税费组成的经费. 在这样的制度下, 各个省均给医院拨划一笔固定的年度预算, 并且医师的工资按论量计酬制发放. 这样的一种经费分配制度就会制约医院的设施 (手术室等) 的可用性, 这反过来也促使医师通过给病人提供过度的服务来增加自己的个人收入. 研究的目标就是要确定一个合理使用手术室的日常时间表. 这个问题的模型是将目标规划和整数规划结合在一起建立的. 第 24 章中的案例 6 对这个问题有详细的研究.

7.1　建立目标规划模型

下面通过一个例题来阐述目标规划的思想.

例 7.1-1 (税收计划)[①]

Fairville 是一个小城市, 拥有大约 2 万居民, 市议会需要制定一张合理的城市税率表. 每年对于实际不动财产的税收基数是 5 亿 5 000 万美元, 对于食品和药品的税收基数是 3 500 万美元, 对于一般销售业的税收基数是 5 500 万美元. 当地每年的汽油消耗预计是 750 万加仑. 市议会希望根据下面 4 个主要的目标确定税率.

(1) 总税款至少要达到 1 600 万美元才能满足市财政承担额.

(2) 食品和药品的税款不能超过总税款的 10%.

(3) 一般销售业的税款不能超过总税款的 20%.

(4) 汽油税不能超过每加仑 2 美分.

用变量 x_p, x_f, x_s 分别表示财产的税率、食品和药品的税率, 以及一般销售业的税率, 用变量 x_g 表示汽油税, 表示每加仑征收 x_g 美分. 那么, 市议会的几个主要目标可以用下面的式子表示 (单位: 100 万美元):

$$\begin{aligned}
550x_p + 35x_f + 55x_s + 0.075x_g &\geqslant 16 && \text{(总税款)}\\
35x_f &\leqslant 0.1(550x_p + 35x_f + 55x_s + 0.075x_g) && \text{(食品/药品税款)}\\
55x_s &\leqslant 0.2(550x_p + 35x_f + 55x_s + 0.075x_g) && \text{(一般销售业的税款)}\\
x_g &\leqslant 2 && \text{(汽油税)}\\
x_p, x_f, x_s, x_g &\geqslant 0
\end{aligned}$$

以上式子可以简化为

$$\begin{aligned}
550x_p + 35x_f + 55x_s + 0.075x_g &\geqslant 16\\
55x_p - 31.5x_f + 5.5x_s + 0.0075x_g &\geqslant 0\\
110x_p + 7x_f - 44x_s + 0.015x_g &\geqslant 0\\
x_g &\leqslant 2\\
x_p, x_f, x_s, x_g &\geqslant 0
\end{aligned}$$

模型中的每一个不等式代表市议会所希望满足的一个目标, 对于这么一些相互矛盾的目标, 也许最好的办法就是寻找一个折中的解.

寻找目标规划的折中解最常用的方法是, 将每一个不等式转换成一个弹性目标. 如果需要的话, 对应的这个弹性目标也可以背离相应的约束. 对于 Fairville 模型而言, 弹性目标可以用下面的式子表示:

$$\begin{aligned}
550x_p + 35x_f + 55x_s + 0.075x_g + s_1^- - s_1^+ &= 16\\
55x_p - 31.5x_f + 5.5x_s + 0.0075x_g + s_2^- - s_2^+ &= 0\\
110x_p + 7x_f - 44x_s + 0.015x_g + s_3^- - s_3^+ &= 0
\end{aligned}$$

① 这个例题依据文献 (Chissman 等人, 1989).

$$x_g + \boldsymbol{s_4^-} - \boldsymbol{s_4^+} = 2$$

$$x_p, x_f, x_s, x_g \geqslant 0$$

$$s_i^-, s_i^+ \geqslant 0, \quad i = 1, 2, 3, 4$$

由于非负变量 s_i^- 和 s_i^+, $i=1,2,3,4$, 代表在第 i 个约束中低于和高于右端项的一种偏离程度, 所以被称为**偏离变量**(deviational variable).

根据偏离变量的定义可以知道, s_i^- 和 s_i^+ 是相关的, 所以它们不能同时作为基变量. 这就意味着在单纯形法的任何一次迭代中每一对偏离变量最多只能有一个被赋值为正数. 对于原来的第 i 个形式为"$\leqslant$"的不等式, 当对应它的 $s_i^- > 0$ 时, 则第 i 个目标是满足的; 当 $s_i^+ > 0$ 时, 那么第 i 个目标是不满足的. 本质上讲, 按照 s_i^- 和 s_i^+ 的定义, 可以根据需要来满足或者背离第 i 个目标. 这就是所谓的在寻找折中解时目标规划的弹性类型. 很自然地, 一个好的折中解就是使得每个目标的背离量尽可能的小.

在 Fairville 模型中, 前 3 个约束都是"$\geqslant$"形式的不等式, 第 4 个约束是"$\leqslant$"形式的不等式, 偏离变量 s_1^-, s_2^-, s_3^-, s_4^+(在模型中用了黑体标出) 代表了对应各自目标的背离量. 那么, 一个折中解就是尽可能满足下面的 4 个目标:

$$\min \quad G_1 = s_1^-$$

$$\min \quad G_2 = s_2^-$$

$$\min \quad G_3 = s_3^-$$

$$\min \quad G_4 = s_4^+$$

此时模型就转变成了在所有等式约束条件下, 求解使得上面这 4 个函数最小.

如何优化一个多目标 (其中甚至包含某些相互矛盾的目标) 的模型呢? 这里提供两种方法求解这种模型: (1) 权和法; (2) 设定优先权法. 7.2 节将详细讲述这两种方法, 它们的共同之处都是将多个目标转化为单个目标.

习题 7.1A

*1. 在 Fairville 税款问题的模型中, 假定城镇政府又指定一项额外的目标 G_5, 要求汽油税不低于总税款的 10%. 再建立相应的目标规划模型.

2. NW 购物商场开展一些特殊的活动来吸引潜在的顾客. 这些活动可以吸引不同年龄段的顾客: 青少年、中年人或者老年人. 两种最流行的活动就是开展乐队音乐会和艺术展览会, 这两种活动每次的花费分别是 1 500 美元和 3 000 美元. 每年给予这两项活动的总经费是 15 000 美元. 商场经理者预计这两项活动的参加人数如下:

活动	每次活动参加的人数		
	青少年	中年人	老年人
乐队音乐会	200	100	0
艺术展览会	0	400	250

经理希望参加的青少年、中年人和老年人的人数至少要分别达到 1 000 人、1 200 人和 800 人. 把这个问题建立成一个目标规划模型.

***3.** 在每个新学年到来的时候, Ozark 大学招生办公室都需要处理许多新生申请. 所有的申请被分为 3 部分: 本州的、外州的和国际的. 本州的申请者和外州的申请者中男女比例分别是 1:1 和 3:2, 国际的学生中男女比例为 8:1. 美国大学测试 (ACT) 的分数是衡量新生是否被接收的一个重要指标. 根据大学的统计数字显示, 本州的、外州的和国际的学生的平均 ACT 分数分别是 27、26 和 23. 对于新生班级, 招生办负责人希望达到下面的一些目标:

(a) 新生班级的人数至少是 1 200 人. **(b)** 所有新生的平均 ACT 分数至少是 25.

(c) 国际的学生至少要占所有新生的 10%. **(d)** 女生与男生的比例至少是 3:4.

(e) 外州的学生至少要占所有新生的 20%.

把这个问题建立成一个目标规划模型.

4. Circle K 农场每天消耗一种特殊的饲料 3 吨. 这种饲料是石灰石、谷物和大豆粗粉的混合物, 它要满足下面所需要的营养:

(a) 钙: 至少要占 0.8%, 但不超过 1.2%. **(b)** 蛋白质: 至少占 22%.

(c) 纤维: 最多占 5%.

下表中给出了这种饲料各个组成原料中的营养成分含量.

原料	每磅组成原料中的营养成分含量		
	钙	蛋白质	纤维
石灰石	0.380	0.00	0.00
谷物	0.001	0.09	0.02
大豆粗粉	0.002	0.50	0.08

将这个问题建立成一个目标规划模型, 针对这种情形谈谈你对有关利用目标规划建立模型的适用范围的认识.

***5.** Mantel 公司生产一种马车玩具, 这种马车在最后装配的时候需要 4 个轮子和 2 个座位. 生产这两种部件的工厂采用每天 3 班生产, 下表给出了 3 班中每个部件的生产数量.

班	每运转一次生产的产品数目	
	轮子	座位
1	500	300
2	600	280
3	640	360

理想情况下, 生产的轮子的数量恰好是座位数量的 2 倍. 然而, 由于生产效率随着班次的不同而不同, 所以恰好满足平衡是不大可能的. Mantel 希望合理安排每个班次中的生产运转次数, 使得生产的 2 种部件的不平衡性尽量小. 由于生产能力的差异, 所以每一个班上都有运转次数的限制: 1 班是 4~5 次, 2 班是 10~20 次, 3 班是 3~5 次. 给这个问题建立一个目标规划模型.

6. Camyo 制造公司生产 4 种零件, 需要用到车床和钻床. 这 2 种机器每天总共运转 10 个小时. 下表给出了每种机器加工每个零件所需要的时间 (分钟):

零件	加工时间	
	车床	钻床
1	5	3
2	6	2
3	4	6
4	7	4

生产商希望平衡 2 台机器的工作时间, 使得这 2 台机器的工作时间差不超过 30 分钟. 市场对每个零件的需求量至少是 10 个单位, 另外零件 1 的数量不能超过零件 2 的数量. 给这个问题建立一个目标规划模型.

7. 生产 2 种产品需要相继使用 2 台机器. 下表给出了 2 种产品每个单位需要的机器加工时间 (分钟):

机器	机器加工时间	
	产品 1	产品 2
1	5	3
2	6	2

每天 2 种产品的生产配额分别是 80 个和 60 个单位, 每台机器每天工作 8 小时, 必要时机器也可以通过超时工作来满足生产的配额. 给这个问题建立一个目标规划模型.

8. Vista 市医院需要提前 4 天制定一个分配剩余床位 (只要当时没有被使用) 来满足病人短暂住院的计划, 在一个 4 天的计划周期内, 大约 30 个、25 个和 20 个病人需要短暂住院的天数分别是 1 天停留、2 天停留和 3 天停留. 在同一个周期内, 剩余床位的数目预计是 20, 30, 30, 30. 建立目标规划来解决当医院允许过量病人住院和不允许过量病人住院这 2 种情况下的问题.

9. 因为 Von Trapp 夫妇都在一个新的城市找到了工作, 所以他们一家计划搬迁到这座城市. 为了选择一个新住房的理想位置, Von Trapp 夫妇列出了他们的下面的几个目标:

(a) 尽量距离 Von Trapp 太太的工作地点比较近 (在 $\frac{1}{4}$ 英里以内).

(b) 尽量远离机场的噪音 (离开机场至少 10 英里).

(c) 要距离商场适当地近 (在 1 英里以内).

Von Trapp 夫妇用城市的地理坐标代表每个候选位置, 并且标记出工作地点、机场和商场的 (x, y) 坐标分别为 (1,1)、(20,15) 和 (4,7), 所有的距离单位是英里, 给这个问题建立一个目标规划模型. [注意: 结果中的约束不是线性的.]

10. *回归分析*. 在一次实验过程中, 假定 y_i 是第 i 个被观察的 (独立的) 因变量的值, $x_{ij}, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$ 是相关自变量的观测值. 希望确定一个线性回归 $y = b_0 + b_1x_1 + \cdots + b_nx_n$ 拟合这些数据, 其中 $b_j, j = 0, 1, \cdots, n$ 表示回归系数. 希望确定所有的 b_j, 使得总的观测值 y_i 和估计值 $b_0 + \sum_{j=1}^{n} b_i x_{ij}$ 之间的绝对离差的总和最小. 给这个问题建立一个目标规划模型.

11. *切比雪夫问题*. 将第 10 题中回归模型的目标变为通过确定 b_j 以最小化最大绝对离差. 给这个问题建立一个目标规划模型.

7.2 求解目标规划的算法

本节将介绍求解目标规划的两种算法, 这两种算法的思想都是用单个目标函数来代替多个目标. 在**权和法** (weights method) 中, 是将问题中代表每个目标的函数进行加权求和以得到单个的目标函数. 在**设定优先权法** (preemptive method) 中, 首先是将所有的目标按照重要性编排一个顺序, 然后模型最优化的过程就是从具有最高优先级的目标开始每次只优化一个目标, 并且使得对较高优先级目标求出的值不会因为求解较低目标时而减少.

因为上面提到的两种方法对于同一个问题经常会给出不同的解, 所以它们是不同的. 然而, 并不能说这两种方法中一个肯定比另外一个好, 因为它们往往适用于不同的决策问题.

7.2.1 权和法

假定一个目标规划模型有 n 个目标, 并且第 i 个目标的形式为

$$\min G_i, \quad i=1,2,\cdots,n$$

权和法中定义的组合目标函数可以表示为

$$\min z = w_1G_1 + w_2G_2 + \cdots + w_nG_n$$

其中参数 $w_i, i=1,2,\cdots,n$ 是正的权重, 反映了决策者对每个目标的相对重要性的一种量化. 例如, 当对于所有的 i, $w_i=1$ 时, 表示所有的目标是同等重要的, 有相同的权重. 然而在确定目标的权重上面是非常主观的, 即便是在文献 (Cohon, 1978) 已经非常成熟的分析程序体系中也带有根深蒂固的主观因素.

例 7.2-1

TopAd 广告代理公司拥有 10 个雇员, 它收到了一份发布一种新产品的合约. 代理商可以通过广播和电视两种途径来发布广告. 下面的表格给出了这两种不同的发布广告的途径可以覆盖的人数以及需要的花费和雇员人数:

	广告的相关数据	
	广播	电视
覆盖人数 (单位: 100 万人)	4	8
花费 (单位: 1 000 美元)	8	24
分配的雇员数目	1	2

合约中要求 TopAd 使用广播广告的时间不超过 6 分钟, 另外广播和电视广告所覆盖的人数需要至少达到 4 500 万人. TopAd 给这个项目的目标预算是 10 万美元. 那么 TopAd 应该如何分配广播和电视的广告时间呢?

令 x_1 和 x_2 分别是分配给广播广告和电视广告的时间, 单位是分钟. 这个问题的目标规划模型可以写成:

$$
\begin{aligned}
&\min\ G_1 = s_1^- \quad \text{(满足覆盖要求)}\\
&\min\ G_2 = s_2^+ \quad \text{(满足预算目标)}\\
&\text{s.t.}\ 4x_1 + 8x_2 + s_1^- - s_1^+ = 45 \quad \text{(覆盖的目标)}\\
&\quad\ 8x_1 + 24x_2 + s_2^- - s_2^+ = 100 \quad \text{(预算的目标)}\\
&\quad\ x_1 + 2x_2 \leqslant 10 \quad \text{(雇员人数限制)}\\
&\quad\ x_1 \leqslant 6 \quad \text{(广播时间限制)}\\
&\quad\ x_1, x_2, s_1^-, s_1^+, s_2^-, s_2^+ \geqslant 0
\end{aligned}
$$

TopAd 公司的经理认为覆盖目标的重要性是资金预算目标重要性的 2 倍, 那么组合在一起的目标函数就是

$$\min\ z = 2G_1 + G_2 = 2s_1^- + s_2^+$$

最优解中, $z = 10$, $x_1 = 5$ 分钟, $x_2 = 2.5$ 分钟, $s_1^- = 5$(即 500 万人), 其他变量都是 0.

从最优值 z 不等于 0 可以看出, 在所要求的目标当中至少有一个是不能够被满足的. 具体来说, $s_1^- = 5$ 意味着对于要求的覆盖目标 (至少覆盖 4 500 万人) 没有达到, 缺了 500 万人. 相反地, 由于 $s_2^+ = 0$, 所以预算目标 (不超过 10 万美元) 是满足的.

目标规划所做的只是给原问题找到一个有效解, 而并非最优的解. 例如, 上面的问题中, 取 $x_1 = 6$ 和 $x_2 = 2$, 同样可以得到相同的覆盖程度 $[4\times 6 + 8\times 2 = 40$ (即 4 000 万人)], 同时花费更低 $[8\times 6 + 24\times 2 = 96$(即 9.6 万美元)]. 实质上, 目标规划是找到一个简单的能够满足模型中目标的解, 而不考虑最优化. 也正是由于这点"缺陷"使得用目标规划来作为一种求解最优化问题的技巧受到置疑.

习题 7.2A

***1.** 考虑习题 7.1A 中处理 Fairville 税款问题的第 1 题. 假定 5 个目标有相同的权重, 求解这个问题. 求出的解是否满足所有的目标呢?

2. 在习题 7.1A 第 2 题中, 假定吸引中年人的目标的重要性是其他两类人 (青少年和老年人) 中任何一类人的重要性的 2 倍. 求出相应的解, 并检查所有的目标是否都得到满足.

3. 在习题 7.1A 第 3 题的 Ozark 大学招生办公室处理新生申请的问题中, 假定新生人数的限制必须满足, 而其他的要求可以看成弹性目标. 另外, 假定 ACT 分数目标的重要性是其他所有目标中任何一个目标重要性的 2 倍.

(a) 求解这个问题, 并检查所有的目标是否都得到满足.

(b) 如果另外假设新生人数的目标也可以看作是弹性目标, 并且这个目标的重要性是 ACT 分数目标重要性的 2 倍, 那么求出的解有何变化.

***4.** 在习题 7.1A 第 4 题的 Circle K 农场模型中, 是否有可能使得所有的营养需求都能够满足?

5. 在习题 7.1A 第 5 题中, 求解这个问题, 并且说明每天生产的轮子数目和座位数目能否达到平衡.

6. 在习题 7.1A 第 6 题中, 假定市场需求目标的重要性是平衡 2 台机器工作时间目标的 2 倍, 同时不允许超时. 求解这个问题, 并且说明所有的目标是否都能够满足.

***7.** 在习题 7.1A 第 7 题中, 假定生产商试图满足对 2 种产品的生产配额要求, 如果需要可以通过额外增加时间来完成. 求出这个问题的解, 并且说明在需要增加时间的情况下, 增加多长时间才可以满足产品配额的要求.

8. 在习题 7.1A 第 8 题的 Vista 市医院问题中, 假定只有床位限制的目标是弹性目标, 并且其他的目标都有相同的重要性. 那么是否所有的目标都可以满足?

9. Malco 公司从其 5 个雇员的档案中汇编出了下面的表格, 来考察影响雇员收入的 3 个因素: 年龄、受教育程度 (表示所就读大学的学制) 和工作年限 (表示参加工作的时间).

年龄	受教育程度 (年)	工作年限	收入 (美元)
30	4	5	40 000
39	5	10	48 000
44	2	14	38 000
48	0	18	36 000
37	3	9	41 000

应用习题 7.1A 第 10 题中目标规划的模型, 使用线性等式 $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$ 来拟合上面的数据.

10. 利用习题 7.1A 第 11 题中的切比雪夫方法求解第 9 题.

7.2.2 设定优先权法

在设定优先权法中, 决策者必须首先对问题中的所有目标做一个重要性的排序. 给定 n 个目标的情形, 问题的目标可以写成

$$\min G_1 = \rho_1 \quad (\text{最高优先权})$$

$$\vdots$$

$$\min G_n = \rho_n \quad (\text{最低优先权})$$

其中 ρ_i 表示第 i 个目标的偏离变量 s_i^- 或者 s_i^+. 例如在例 7.2-1 的 TopAd 模型中, $\rho_1 = s_1^-$, $\rho_2 = s_2^+$.

求解的过程是从具有最高优先权的目标 G_1 开始, 每次只考虑一个目标, 直到考虑完具有最低优先权的目标 G_n 为止. *在这个过程中, 求解较低优先权的目标时得到的解不会降低已经求出的较高优先权目标的解的函数值.*

目标规划的相关文献已经给出了一种可以确保不会降低具有较高优先权目标的“特殊”单纯形法. 这种方法采用的是**列删除规则** (column-dropping rule), 列删除规则是指在求解目标 G_{k+1} 之前, 将求解目标 G_k 得到的最优单纯形表中具有非零的简约费用 $(z_j - c_j \neq 0)$ 的非基变量 x_j 删除. 这是由于规则中所选中的那些非基变量在接下来的目标优化中, 如果取值大于零, 就会降低 (当然永远不会改进) 对于以前求出的较高优先权目标的解的函数值. 为了方便操作, 会对单纯形表的样式进行一些改变, 使得它可以包含模型中所有目标的目标函数.

上面提出的列删除不会增加目标规划的复杂性. 在这里, 我们将证明可以通过下面步骤中更加直接的方法得到上面的结果.

第 0 步　确定模型中的目标, 并且按优先权排序:

$$G_1 = \rho_1 \succ G_2 = \rho_2 \succ \cdots \succ G_n = \rho_n$$

令 $i = 1$.

第 i 步　求解最小化 G_i 的线性规划问题 LP_i, 并且令 $\rho_i = \rho_i^*$ 来定义相应最优值的偏离变量 ρ_i. 如果 $i = n$, 停止; LP_n 解决了 n 目标规划. 否则, 在问题 G_i 上增加约束 $\rho_i = \rho_i^*$ 以确保在以后的问题中 ρ_i 的值不会降低. 令 $i = i + 1$, 重复执行第 i 步.

虽然相继增加的特殊约束 $\rho_i = \rho_i^*$ 不像列删除规则理论上那样整齐, 但是它也恰好可以得到同样的效果, 并且更重要的是非常容易理解.

从计算复杂性的角度来说, 列删除规则更具有优势. 实质上, 列删除规则通过删除变量使得问题规模一步步地变小, 而我们上面的步骤中确会因为约束的增多而变得复杂. 然而, 考虑到所增加约束 $(\rho_i = \rho_i^*)$ 的特点, 我们可以通过直接替换变量 ρ_i 来修改单纯形算法, 以满足所增加的约束. 这种替换只会影响含有 ρ_i 的约束, 并且在求解下一个目标时减少了变量个数. 当然我们还可以应用 7.4.2 节中提到的有界的单纯形法, 就是将约束 $\rho_i = \rho_i^*$ 换成 $\rho_i \leqslant \rho_i^*$, 在这种情况下增加的约束就成了隐式的. 考虑到以上因素, 列删除规则在理论上不错, 但是实质上从计算复杂性的角度也没有特别的优势. 为了文章的完整性, 我们还将通过例 7.2-3 来详细描述列删除规则的计算过程.

例 7.2-2

下面用设定优先权法求解例 7.2-1. 假定覆盖的目标具有较高的优先权.

第 0 步　$G_1 \succ G_2$

$$G_1 : \min\ s_1^- \quad (\text{满足覆盖的目标})$$

$$G_2 : \min\ s_2^+ \quad (\text{满足预算的目标})$$

第 1 步　求解线性规划 LP_1:

$$
\begin{aligned}
\min\ & G_1 = s_1^- \\
\text{s.t.}\ & 4x_1 + 8x_2 + s_1^- - s_1^+ = 45 \quad (\text{覆盖的目标}) \\
& 8x_1 + 24x_2 + s_2^- - s_2^+ = 100 \quad (\text{预算的目标}) \\
& x_1 + 2x_2 \leqslant 10 \quad (\text{雇员人数限制}) \\
& x_1 \leqslant 6 \quad (\text{广播时间限制}) \\
& x_1, x_2, s_1^-, s_1^+, s_2^-, s_2^+ \geqslant 0
\end{aligned}
$$

根据 TORA 得到最优解是 $x_1 = 5$ 分钟, $x_2 = 2.5$ 分钟, $s_1^- = 5$(500 万人), 其他变量为零, 从这个解可以看出对于覆盖目标 G_1, 缺少 500 万人.

在 LP_1 中, 我们有 $\rho_1 = s_1^-$, 所以对于第二个目标的问题 G_2 增加约束 $s_1^- = 5$(或者 $s_1^- \leqslant 5$).

第 2 步 下面需要求解 LP_2, 目标函数是

$$\min\ G_2 = s_2^+$$

约束集合除了包含第 1 步中的约束外, 还有增加的约束 $s_1^- = 5$. (在 TORA 的 MODIFY 选项中可以很容易地将新增加的约束替代上, 只需要将变量 s_1^+ 的上下界都设为 5.)

因为求解 LP_1 得到的最优解中 $s_2^+ = 0$, 所以不需要再求解 LP_2, LP_1 的解自动地优化了 LP_2. 由 $s_2^+ = 0$ 可以知道目标 G_2 是满足的.

由于增加的约束是 $s_1^- = 5$, 所以我们可以将第一个约束中的 s_1^- 进行替代, 那么这个约束右边代表覆盖目标的值就从 45 降到 40, 因此可以将 LP_2 简化为

$$
\begin{aligned}
\min\ & G_2 = s_2^+ \\
\text{s.t.}\ & 4x_1 + 8x_2 - s_1^+ = 40 \quad (\text{覆盖的目标}) \\
& 8x_1 + 24x_2 + s_2^- - s_2^+ = 100 \quad (\text{预算的目标}) \\
& x_1 + 2x_2 \leqslant 10 \quad (\text{雇员人数限制}) \\
& x_1 \leqslant 6 \quad (\text{广播时间限制}) \\
& x_1, x_2, s_1^+, s_2^-, s_2^+ \geqslant 0
\end{aligned}
$$

从这个新的规划中可以看出, 变量个数比 LP_1 少了一个, 这也恰好反映出了*列删除规则*的思想.

例 7.2-3 (列删除规则)

在这个例子中, 我们将证明对于例 7.2-1 和例 7.2-2 中的问题, 如果采用设定优先权法来*最优化*目标而不是仅仅*满足*目标, 可以得到一个更好的解. 在这个例题中, 我们还将演示, 如何采用*列删除规则*求解目标规划.

将例 7.2-1 中的目标重新叙述为

$$\text{优先权1}: \text{最大化覆盖}(P_1)$$
$$\text{优先权2}: \text{最小化花费}(P_2)$$

这两个目标可以写成数学表达式就是

$$\begin{aligned} \max\ P_1 &= 4x_1 + 8x_2 \quad (\text{覆盖}) \\ \min\ P_2 &= 8x_1 + 24x_2 \quad (\text{花费}) \end{aligned}$$

例 7.2-1 和例 7.2-2 中对覆盖和花费目标的限制 (分别为 45 和 100) 就不需要了, 因为我们将使用单纯形法使得这些限制最优化.

因此新的问题可以描述为

$$\begin{aligned} \max\ P_1 &= 4x_1 + 8x_2 \\ \min\ P_2 &= 8x_1 + 24x_2 \\ \text{s.t.}\quad x_1 + 2x_2 &\leqslant 10 \\ x_1 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$

首先使用例 7.2-2 中的过程求解这个问题.

第 1 步 求解线性规划 LP_1:

$$\begin{aligned} \max\ P_1 &= 4x_1 + 8x_2 \\ \text{s.t.}\quad x_1 + 2x_2 &\leqslant 10 \\ x_1 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$

根据 TORA 得到最优解是 $x_1 = 0$ 分钟, $x_2 = 5$ 分钟, $P_1 = 40$, 说明最大可以覆盖的人数是 4 000 万人.

第 2 步 增加约束 $4x_1 + 8x_2 \geqslant 40$, 以保证在下面的求解过程中不会使得目标 G_1 的值降低. 因此可以得到 LP_2, 新的问题可以描述为

$$\begin{aligned} \min\ P_2 &= 8x_1 + 24x_2 \\ \text{s.t.}\quad x_1 + 2x_2 &\leqslant 10 \\ x_1 &\leqslant 6 \\ 4x_1 + 8x_2 &\geqslant 40 \quad (\text{增加的约束}) \\ x_1, x_2 &\geqslant 0 \end{aligned}$$

求解 LP_2 可以得到, $P_2 = 96$(即 9.6 万美元), $x_1 = 6$ 分钟, $x_2 = 2$ 分钟. 这个解同样也可以保证覆盖的人数是 4 000 万, 而且这个解对应的花费要比例 7.2-2 中求出的解所对应的花费要小, 这是因为上一个例题所要求的只是满足目标, 而不是最优化目标.

下面我们使用**列删除规则**来求解这个问题. 在单纯形表中, 我们将所有的目标都放在其中分别作为行.

LP_1(覆盖最大化) 在线性规划 LP_1 的单纯形表中, 同时包含 P_1 和 P_2 目标行. 第一步只对 P_1 目标行进行优化, 这时的 P_2 目标行是不起作用的, 但是为了方便以后求解 LP_2, 同时也需要对 P_2 目标行进行修改.

求解 LP_1 需要两步迭代:

迭代	基	x_1	x_2	s_1	s_2	解
1	P_1	-4	-8	0	0	0
	P_2	-8	-24	0	0	0
	s_1	1	2	1	0	10
	s_2	1	0	0	1	6
2	P_1	0	0	4	0	40
	P_2	4	0	12	0	120
	x_2	$\frac{1}{2}$	1	$\frac{1}{2}$	0	5
	s_2	1	0	0	1	6

最后一张单纯形表给出了最优解 $x_1 = 0, x_2 = 5, P_1 = 40$.

使用**列删除规则**是在进入求解 LP_2 之前, 从 LP_1 的最优单纯形表中删除满足 $z_j - c_j \neq 0$ 的**非基变量** x_j. 这样做的原因就是, 如果保留像这样的一些变量, 在以后求解较低优先权目标的过程中就会取值增加, 从而降低求解较高优先权目标的解的质量.

LP_2(费用最小化) 列删除规则将变量 $s_1(z_j - c_j = 4)$ 从单纯形表中删除. 如果 s_1 不被删除, 我们看 P_2 所在的行, 在接下来开始的第一次迭代中, s_1 就将成为进基变量, 从而求出的最优解是 $x_1 = x_2 = 0$, 显然这会使得目标 P_1 的值从 $P_1 = 40$ 降低到 $P_1 = 0$. (读者可以试一试!)

目标 P_2 是最小化类型. 首先删除了变量 s_1, 然后选择进基变量, 由于对应 x_1 的 $z_j - c_j = 4(\geqslant 0)$, 所以选择 x_1 可以改进 P_2 的目标值. 下面的表格为我们展示了迭代的过程. 由于我们不会再考虑 LP_1 对应的目标, 所以将 P_1 对应的行删除了.

迭代	基	x_1	x_2	s_1	s_2	解
1	P_1					40
	P_2	4	0		0	120
	x_2	$\frac{1}{2}$	1		0	5
	s_2	1	0		1	6
2	P_1					40
	P_2	0	0		-4	96
	x_2	0	1		$-\frac{1}{2}$	2
	x_1	1	0		1	6

最优解是 $x_1 = 6, x_2 = 2$, 相应的覆盖人数是 $P_1 = 40$, 总的花费是 $P_2 = 96$. 这与前面求出的结果是一致的.

AMPL 程序

AMPL 可以很容易地应用例 7.2-2 中给出的思想来进行求解, 也就是增加约束来保证已经求出的较高优先权目标的值在以后的求解过程中不会降低. 图 7.1 给出了一般 AMPL 的代码, 可以交互式地应用设定优先权法.(见文件 amplEx7.1-1.txt)

```
param n;
param m;
param r;
param p;
param q;
param a{1..m,1..n};
param b{1..m};
param um{1..m} default 100000;
param up{1..m} default 100000;
var x{1..n} >=0;
var sminus{i in 1..r}>=0,<=um[i];
var splus{i in 1..r}>=0,<=up[i];

minimize z{i in 1..r}: p*sminus[i]+q*splus[i];
subject to
c1{i in 1..r}:sum{j in 1..n}a[i,j]*x[j]+sminus[i]-splus[i]=b[i];
c2{i in r+1..m}: sum{j in 1..n} a[i,j]*x[j]<=b[i];
data;
param m:=4;
param n:=4;
param r:=4;
param a: 1      2       3      4:=
     1  550    35      55     .075
     2  55     -31.5   5.5    .0075
     3  110    7       -44    .015
     4  0      0       0      1;
param b:=1 16 2 0   3 0  4 2;
```

图 7.1 交互式应用设定优先权法的 AMPL 模型 (文件 amplEx7.1-1.txt)

除了便于交互式地应用设定优先权法的接口程序外, AMPL 模型的设计是非常标准的. 具体来说, 模型设定了前 r 个约束是目标约束, 剩下的 $(m-r-1)$ 个是问题固有的精确约束. 模型可以包含 r 个不同的目标函数, 通过下面的索引 AMPL 语句可以使得这些目标函数包含在同一个模型中 (只有提前编入索引的名字才能作为目标函数).

```
minimize z{i in 1..r}:p*sminus[i]+q*splus[i];
```

根据这样定义的目标函数, 可以通过设定 $(p=1,q=0)$ 和 $(p=0,q=1)$, 分别达到最小化 $z_i=s_i^-$ 和 $z_i=s_i^+$ 的目标.

在求解的过程中每次从一个优先级到下一个优先级时, 都需要增加新的约束, AMPL 不需要这么做, 它利用一个规划的技巧, 通过改变偏离变量的上界来实现. 用参数 um[i] 和 up[i] 分别表示 sminus[i](s_i^-) 和 splus[i](s_i^+) 的上界. 可以通过修正这两个参数来满足类型 sminus[i](s_i^-) <= um[i] 和 splus[i](s_i^+) <= up[i] 的隐式约束. 优先权为 i 的目标对应的 um 和 up 是由前 $(i-1)$ 个目标求出的解确定. 其中参数 um 和 up 的初始 (默认) 值设为无穷大.

下面将简短地说明 AMPL 模型是如何确定r个目标函数, 设定p和q的值, 以及给出 s_i^- 和 s_i^+ 的上界. 这些都是交互式的, 使得 AMPL 模型可以更便利地求解目标规划.

使用例 7.1-1 中的数据, 模型的目标是

$$\min \quad G_1=s_1^-$$
$$\min \quad G_2=s_2^-$$
$$\min \quad G_3=s_3^-$$
$$\min \quad G_4=s_4^+$$

假定目标的优先权顺序是

$$G_2 \succ G_1 \succ G_3 \succ G_4$$

那么 AMPL 模型的运行过程按照下面进行: 对 G_2, 因为我们是最小化 z[2] =sminus[2], 所以令 $p=1$ 和 $q=0$. 利用下面的一组命令来进行运算:

```
ampl: model amplEx7.1-1.txt;
ampl: let p:=1;let q:=0;
ampl: objective z[2];
ampl: solve; display z[2], x, sminus, splus;
```

执行这些命令输出的结果是

```
z[2] = 0
:        x            sminus splus :=
1  2.62361e-28       16     0
2  0                  0     0
3  8.2564e-28         0     0
4  0                  2     0
```

这个解 $(x_1=0,\ x_2=0,\ x_3=0,\ x_4=0)$ 说明由于z[2]=0(也就是 $s_2^-=0$) 所以目标 G_2 是满足的, 但是对于第二优先权 G_1 是不满足的 (因为 $s_1^-=16$), 因此我

们需要在不会影响目标 G_2 的前提下, 继续优化目标 G_1, 这个过程首先需要根据目标 G_2 的解来改变变量 s_2^- 的上界, 也就是变成 0. 对于目标 G_1, 由于所求的目标是最小化 s_1^-, 所以此时的 $p=1$ 和 $q=0$ 不发生变化. 这些可以使用下面的交互式 AMPL 命令实现:

```
ampl: let um[2]:=0;
ampl: objective z[1]; solve; display z[1],x, sminus, splus;
```

输出的结果是

```
z[1] = 0
:   x           sminus splus :=
1   0.0203636   0      0
2   0.0457143   0      0
3   0.0581818   0      0
4   0           2      0
```

这个结果告诉我们, 剩余的目标都可以满足, 因此不需要再继续优化以后的 2 个目标. 目标规划的解就是, $x_p = 0.020\ 363\ 6$, $x_f = 0.045\ 714\ 3$, $x_s = 0.058\ 181\ 8$ 和 $x_g = 0$.

评注 (1) 如果用语句 `fix sminus[2]:=0;` 或者 `let sminus[2]:=0` 来代替语句 `let um[2]:=0;`, 可以得到相同的结果.

(2) 在交互式语句中, 可以通过一个 commands 文件来自动地选择当前优化的目标, 并且可以在进入下一个目标求解之前自动地加上准确的限制约束, 这个文件命名为 amplCmds.txt, 使用这个文件需要对原始的文件 amplEx7.1-1A.txt 中的模型进行一些修改. 为了实现的完整性, 将模型的数据单独的存放在文件 amplData.txt 中. 在这种情况下, 模型的执行需要下面 3 个命令行:

```
ampl: model amplEx7.1-1A.txt;
ampl: data amplData.txt;
ampl: commands amplCmds.txt;
```

需要了解更多 commands 的用法可以参看附录 A.7.

习题 7.2B[①]

1. 在例 7.2-2 中, 假定预算目标增加到 11 万美元, 覆盖人数的目标不变仍然是 4 500 万人. 利用设定优先权法求解这个问题.

2. 利用下面设定的优先权顺序, 求解习题 7.1A 中的第 1 题.

$$G_1 \succ G_2 \succ G_3 \succ G_4 \succ G_5$$

① 使用交互式的 AMPL 模型求解这些问题会更加方便.

3. 考虑习题 7.1A 第 2 题中关于 NW 购物商场确定乐队音乐会和艺术展览会的问题. 用 G_1, G_2, G_3 分别代表吸引青少年、中年人和老年人的目标. 分别按照下面的优先权顺序求解这个目标规划问题:

(a) $G_1 \succ G_2 \succ G_3$ (b) $G_3 \succ G_2 \succ G_1$

说明目标满足与否取决于优先权顺序.

4. 利用设定优先权法求解习题 7.1A 第 3 题的 Ozark 大学模型, 假定目标的优先权顺序与问题给出的顺序一致.

参 考 文 献

Chissman, J., T. Fey, G. Reeves, H. Lewis, and R. Weinstein, "A Multiobjective Linear Programming Methodology for Public Sector Tax Planning," *Interfaces*, Vol. 19, No. 5, pp. 13-22, 1989.

Cohon, T. L., *Multiobjective Programming and Planning*, Academic Press, New York, 1978.

Ignizio, J. P., and T. M. Cavalier, *Linear Programming*, Prentice Hall, Upper Saddle River, NJ, 1994.

Steuer, R. E., *Multiple Criteria Optimization: Theory, Computations, and Application*, Wiley, New York, 1986.

第 8 章　整数线性规划

本章导读　整数线性规划 (Integer Linear Program, ILP) 是要求某些或者全部的变量只取整数值 (或者离散值) 的线性规划. 在学习整数线性规划的过程中, 需要集中在以下 3 点: 应用、理论和计算. 为了说明整数线性规划在实际中的广泛应用, 本章首先介绍大量相关的应用实例. 然后给出整数线性规划的两种经典算法: 分支限界法 (Branch and Bound, B&B) 和割平面法. 在这两种方法中, 分支限界法计算更加有效, 实际上, 几乎全部的商用软件都是基于分支限界法的思想. 本章的最后介绍一个有着非常重要的实际应用的问题: 旅行商问题 (Traveling Salesperson Problem, TSP).

整数线性规划算法的一个缺点就是, 在解决整数问题缺乏一致性. 虽然已经从理论上证明这些算法都是可以在有限步内收敛的, 但是它们在计算机上实现起来是很不一样的 (由于计算机固有的误差舍入方式不同). 因此在学习本章的过程中, 读者要注意这一点.

本章还将演示 AMPL 和 Excel 规划求解是如何应用到求解整数线性规划问题的. 在 TORA 的用户使用选项中, 将可以看到分支限界法的详细计算过程.

本章包括 1 个实际应用的综述、12 个例题、5 个 AMPL 模型、1 个 Excel 电子表格、65 个节后习题和 10 个案例. 所有的案例都放在本书下册附录 E 中, AMPL/Excel/Solver/TORA 的程序放在网络资源文件夹 ch8Files 中.

实际应用——PFG 优化拖车的载荷问题

PFG 使用一种特殊的 5 轮拖车给顾客运送成捆的平板玻璃. 捆的大小和重量可以不同, 根据收到订单的不同, 一辆拖车可以装载不同型号的捆. 政府规定了车轴可以承受的最大重量, 一辆拖车可以装载多少捆取决于这些捆放在拖车上的位置. 问题就是在不超过政府规定的车轴承载能力的限制条件下, 如何使得拖车装载更多的捆. 这个问题可以转化为整数规划求解. 第 24 章中的案例 7 对这个问题有详细的描述和研究.

8.1　应 用 实 例

本节将给出一些整数线性规划的应用实例, 这些应用的实例一般可以分为两类: 直接的和转化的. 在直接的应用类中, 变量自然的就是整数型的, 并且可以假

定是二元的 (0 或者 1) 或者一般离散型的. 例如, 某个问题会涉及确定一项方案是否被选择执行 (二元的), 或者确定完成某项任务需要的最优机器数目 (一般离散值). 在*转化的*的应用类中, 原始问题可能不涉及整数型变量, 这是难以分析处理的, 但可以通过使用辅助整数型变量 (一般是二元的) 将问题转化成容易处理的. 例如, 在一台机器上安排两项工作 A 和 B 的顺序时, 工作 A 可以在工作 B 之前加工或者工作 B 在工作 A 之前加工. 因为所有的数学规划算法都是仅处理约束之间是"并"的情形, 而这里对于约束要求"或"的关系使得问题变得很难分析处理. 对于这样的情况, 通过引入辅助二元变量将"或"的约束转化成*等价的*"并"的约束.

为方便起见, 当所有变量都是整数型时称为**纯** (pure) 整数规划; 否则, 如果一个问题中既有连续的变量又有整数的变量, 那么就称为**混合** (mixed) 整数规划.

8.1.1 资本预算

本节将研究在个人项目中决定是否需要进行某些投资. 既要考虑到这些投资在个人项目中带来的收益, 又要考虑总的预算是有限制的.

例 8.1-1 项目选择

在一个 3 年的规划周期内, 有 5 个项目可供选择. 下面的表格给出了每一个项目可以带来的期望收益以及相应每年的支出 (单位均为 100 万美元).

项目	每年支出			收益
	1	2	3	
1	5	1	8	20
2	4	7	10	40
3	3	9	2	20
4	7	4	1	15
5	8	6	10	30
可用资金	25	25	25	

那么在这个 3 年规划周期应该选择哪些项目呢?

问题可以化成一个对于每个项目的选择为"是–否"的决策, 引入二元变量 x_j,

$$x_j = \begin{cases} 1, & \text{如果选择项目 } j \\ 0, & \text{如果不选择项目 } j \end{cases}$$

那么整数线性规划模型是

$$\begin{aligned} \max\ z = 20x_1 + 40x_2 + 20x_3 + 15x_4 + 30x_5 \\ \text{s.t.}\ 5x_1 + 4x_2 + 3x_3 + 7x_4 + 8x_5 \leqslant 25 \\ x_1 + 7x_2 + 9x_3 + 4x_4 + 6x_5 \leqslant 25 \\ 8x_1 + 10x_2 + 2x_3 + x_4 + 10x_5 \leqslant 25 \\ x_1, x_2, x_3, x_4, x_5 = (0,1) \end{aligned}$$

最优整数解 (可以使用 AMPL, Excel 规划求解或者 TORA 求解)[1] 是 $x_1 = x_2 = x_3 = x_4 = 1$, $x_5 = 0$, 对应的最优值是 $z = 95$. 从这个解可以看出, 除了第 5 个项目, 其他所有项目都要选择.

评注 如果将连续的线性规划问题的解与整数线性规划的解进行比较, 可以得到一些有趣的结论. 首先将上面例题中的整数约束要求 $x_j = (0,1)$ 替换为 $0 \leqslant x_j \leqslant 1$, 那么求出的最优解是: $x_1 = 0.578\,9$, $x_2 = x_3 = x_4 = 1$, $x_5 = 0.736\,8$, 对应的最优值是 $z = 108.68$. 当然这个解因为有两个变量取值为小数, 所以在实际中是不合理的, 如果我们采用舍入方法, 将得到的解变成离它最近的整数, 那么就得到 $x_1 = x_5 = 1$, 这意味着 5 个项目都要选择, 显然是一个不可行的解. 更重要的是, 因为这个问题中变量 x_j 代表的是一个"是–否"的决策, 所以舍入法不是可行的.

习题 8.1A[2]

1. 根据下面的约束条件, 将例 8.1-1 中的资本预算模型进行修改并求解:
(a) 如果选择了项目 1 或者项目 3, 那么一定要选择项目 5.
(b) 项目 2 和项目 3 相互排斥.

2. 需要将一个容器放满五类物品, 物品 i 的重量、体积和单价分别用 w_i, v_i, r_i 表示, 下表中给出了相应的数值.

物品 i	单位重量 w_i(吨)	单位体积 v_i(yd^3)	单价 r_i(100 美元)
1	5	1	4
2	8	8	7
3	3	6	6
4	2	5	5
5	7	4	4

最大可允许装的物品重量和体积分别是 112 吨和 109 yd^3. 建立整数线性规划模型, 并使装入物品的价值最大.

***3.** 假定有 7 个装满酒的瓶子、7 个装了半瓶酒的瓶子和 7 个空的瓶子. 现在需要将这 21 个瓶子分给 3 个人, 使得每个人都分得 7 个瓶子. 另外, 要保证每个人都得到等量的酒. 将这

---

① 使用 TORA 时, 在 Main Menu 中选择 Integer Programming. 输入问题数据之后, 跳出输出窗口并运行 Automated B&B 得到最优解. Excel 规划求解的用法除了需要声明变量是整数之外, 与求解线性规划问题的过程相同. 当增加一个新约束的时候, **规划求解参数**对话框中的整数型选项 (*int* 或者 *bin*) 是可选用的. 使用 AMPL 求解整数规划的过程, 除了需要在目标变量的定义声明语句中通过添加关键词`integer`(或者`binary`) 来声明某些或者全部变量为整数之外, 其余一切都与求解线性规划的过程相同. 例如, 变量声明语句`var x{J}>=0, integer;` 就声明了对于所有的 $x_j (j \in J)$ 都是非负整数型变量. 如果 x_j 是二元的, 声明语句应为`var x{J} binary;`. 执行求解语句`solve;`, 需要加上声明语句`option solver cplex;`.

② 第 3 题至第 6 题来源于 Malba Tahan, *El Hombre que Calculaba*, Editorial Limusa, Mexico City, pp.39-182, 1994.

个问题描述成整数线性规划约束, 并且求解. (*提示*: 引入一个虚拟的目标函数, 这个函数的所有系数都为 0.)

4. 一个古怪的阿拉伯酋长留下了一份遗嘱, 遗嘱中将他的骆驼群分给他的 3 个儿子. Tarek 至少得到半数的驼群, Sharif 至少得到驼群的 $\frac{1}{3}$, Maisa 至少得到驼群的 $\frac{1}{9}$, 剩余的捐献给慈善机构. 遗嘱中并没有指出到底驼群的数目是多少, 只是告诉了这个骆驼群的数目是个奇数, 并且这个指定的慈善机构恰好得到了一匹骆驼. 利用整数线性规划模型确定这个酋长到底留下了多少匹骆驼, 并且指出每个儿子各得到多少匹.

5. 一对农场夫妇让他们的 3 个儿子带着 90 个苹果到市场上去卖, 并且分别交给他们不同数目的苹果以及告诉了相应出售的价格. 最大的儿子 Karen 拿到 50 个苹果, 较小的儿子 Bill 拿到 30 个苹果, 最小的儿子 John 只拿到 10 个苹果. 这对夫妇规定了 5 条原则: (a) 出售的价格或者是 1 美元 7 个苹果或者是 3 美元 1 个苹果, 或者两种价格组合在一起; (b) 每个孩子都可以选择价格中的一种或者两种; (c) 每个孩子出售苹果得到的钱数相同; (d) 每个孩子出售得到的钱数都是整数美元 (不能有美分); (e) 在所有规定的条件下, 每个孩子出售得到的钱数尽可能多. 假定每个孩子都能够卖光他们自己携带的苹果, 利用整数线性规划模型确定这 3 个孩子是如何满足他们父母要求的.

***6.** 从前, 有一个商船的船长想要奖励他的 3 个船员, 因为他们在大海上暴风雨中奋不顾身抢救船上货物. 这名船长从会计办公室拿出了一定数目的钱, 并且让第一主管在船靠岸之后将这些钱平均分给这 3 名勇敢的船员. 一天夜里, 其中一名船员悄悄地来到了主管的办公室, 并要求提前得到这些钱的 $\frac{1}{3}$. 当他将钱分成 3 份之后还剩余 1 枚金币, 于是这名船员拿走其中的一份和这剩余的 1 枚金币. 第二天晚上, 另外一名船员也悄悄来到主管的办公室, 他要求做同样的事情, 于是他将剩余的钱等分为 3 份, 最后又剩余 1 枚金币, 这名船员同样拿走了其中的一份和这枚剩余的金币. 第三天晚上, 最后一名船员也做了相同的事情, 带走了剩余钱币等分成的 3 份中的 1 份以及不可等分而剩余的 1 枚金币. 船靠岸之后, 主管将剩余的金币平均分成 3 份, 还是剩余 1 枚金币, 于是主管决定保留这枚金币, 将其他金币分给这 3 名船员. 请问最初一共有多少枚金币? 将这个问题转化成整数线性规划模型, 并求解. (*提示*: 这个问题有无穷多个整数解. 为简单起见, 求满足条件的钱数最小的整数解. 然后在这个最小整数解的基础上加 1, 以它作为钱数的下界, 可以求解得到问题次小的整数解, 依此类推可以得到这个问题一般解的形式.)

7. (Weber, 1990) 有下面一些 3 个字母的单词: AFT、FAR、TVA、ADV、JOE、FIN、OSF 和 KEN. 假定我们给字母表中的 26 个字母赋值, 从 $A = 1$ 到 $Z = 26$. 上面的每个单词的得数就是组成它的 3 个字母的数值的和, 例如, AFT 的得数是 $1 + 6 + 20 = 27$. 从上面的 8 个单词中选择 5 个单词, 使得总的得数最大. 同时所选择的 5 个单词要满足下面的条件:

$$\begin{pmatrix}\text{第 1 个字母位上}\\ \text{的得数和}\end{pmatrix} < \begin{pmatrix}\text{第 2 个字母位上}\\ \text{的得数和}\end{pmatrix} < \begin{pmatrix}\text{第 3 个字母位上}\\ \text{的得数和}\end{pmatrix}$$

将这个问题转化成整数线性规划模型并求解.

8. 对于第 7 题, 除了要求总的得数最大, 还要求第 1 列的总得数和第 2 列的总得数也最大. 求出此时的最优解.

9. (Weber, 1990) 考虑下面的两组单词:

第 1 组	第 2 组
AREA	ERST
FORT	FOOT
HOPE	HEAT
SPAR	PAST
THAT	PROF
TREE	STOP

两组中的所有单词都是从下面 9 个字母取出的: A, E, F, H, O, P, R, S, T. 建立一个模型, 给这 9 个字母从 1 到 9 赋值, 使得这两组总得数的差值最小. [注: 一个单词的得数是组成它的几个字母的值的和.]

***10.** Record-a-Song 公司与一位新的歌星签约录制 8 首歌曲, 这 8 首歌曲的时间长度分别是 8, 3, 5, 5, 9, 6, 7, 12 分钟. Record-a-Song 公司打算使用一种双面的录音磁带进行录音, 每一面的容量为 30 分钟. 公司希望将所有的歌曲分配在磁带的两面, 使得两面的歌曲时间总长度尽量相同. 根据这个问题, 建立整数线性规划模型, 并求出最优解.

11. 在第 10 题中, 假定由于特定的音质要求第 3 首和第 4 首歌曲不能录制在磁带的同一面, 将这个问题转化成整数线性规划模型. 请问是否可以在一个每面长度为 25 分钟的磁带上录制这 8 首歌曲? 如果不能, 应用整数线性规划来确定录制这 8 首歌曲所需的磁带容量最小.

***12.** (Graves 等人, 1993) 考虑到学校资源和教室的限制, Ulern 大学采用数学模型来优化学生对课程的满意度. 为了演示这个模型的应用, 我们使用一个简化的例子, 假定由 10 个学生需要从学校设立的 6 门课程中选择 2 门. 下表中给出了每一个学生对每一门课程的满意度评分, 其中分值 100 是最高满意度. 为简便起见, 假定对于所选择的两门课程的总满意度评分是这两门课程满意度评分的和. 课程容量代表了这门课程最多能接纳的学生数目.

学生	课程满意度评分					
	1	2	3	4	5	6
1	20	40	50	30	90	100
2	90	100	80	70	10	40
3	25	40	30	80	95	90
4	80	50	60	80	30	40
5	75	60	90	100	50	40
6	60	40	90	10	80	80
7	45	40	70	60	55	60
8	30	100	40	70	90	55
9	80	60	100	70	65	80
10	40	60	80	100	90	10
课程容量	6	8	5	5	6	5

根据这个问题, 建立整数线性规划模型, 并求出最优解.

8.1.2　集合覆盖问题

在这一类问题中, 会有许多服务装置为一些设备提供互相重叠的服务, 目标就

是要确定安装数目最少的装置来覆盖每一个设备 (满足服务需求). 例如, 几个污水处理工厂可以选择建造在几个不同的位置, 在不同位置可以服务不同的几个城市, 当一个城市可以得到几个不同的工厂服务的时候就是重叠服务.

例 8.1-2 安装安全专用电话

为了提高校园的安全性, A 大学的保安部门决定在校园内部的几个位置安装紧急报警电话. 保安部希望在校园的每条主要街道上都至少有一部电话的情况下, 使得安装的总电话数目最少. 图 8.1 给出了校园的主要街道图 (A 到 K).

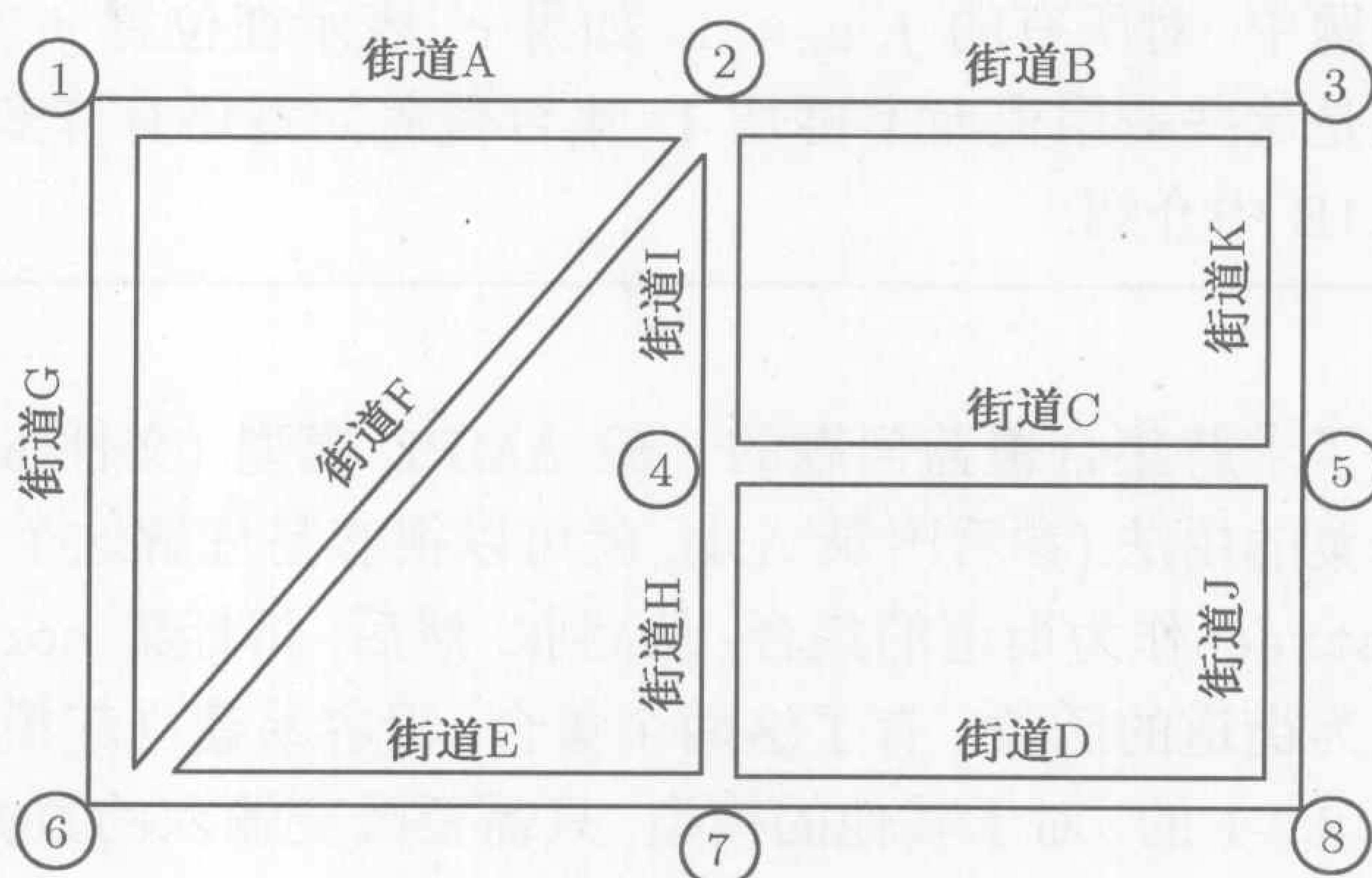


图 8.1 A 大学校园的街道地图

将电话安装在街道的交叉口处是比较合理的, 因为这样就可以至少为两条街道提供服务. 按照图 8.1 中街道的设计可以看出, 最多需要安装 8 部电话.

定义

$$x_j = \begin{cases} 1, & \text{如果位置 } j \text{ 安装电话} \\ 0, & \text{如果位置 } j \text{ 不安装电话} \end{cases}$$

问题是求每一条街道都至少安装一部电话. 那么, 模型可以写成:

$$\begin{aligned}
\min\ & z = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 \\
\text{s.t.}\ & x_1 + x_2 \geqslant 1 \quad (\text{街道 A}) \\
& x_2 + x_3 \geqslant 1 \quad (\text{街道 B}) \\
& x_4 + x_5 \geqslant 1 \quad (\text{街道 C}) \\
& x_7 + x_8 \geqslant 1 \quad (\text{街道 D}) \\
& x_6 + x_7 \geqslant 1 \quad (\text{街道 E}) \\
& x_2 + x_6 \geqslant 1 \quad (\text{街道 F}) \\
& x_1 + x_6 \geqslant 1 \quad (\text{街道 G}) \\
& x_4 + x_7 \geqslant 1 \quad (\text{街道 H}) \\
& x_2 + x_4 \geqslant 1 \quad (\text{街道 I})
\end{aligned}$$

$$x_5 \quad + x_8 \geqslant 1 \text{ (街道 J)}$$

$$x_3 \quad + x_5 \quad \geqslant 1 \text{ (街道 K)}$$

$$x_j = (0,1), \quad j = 1, 2, \cdots, 8$$

这个问题的最优解需要安装 4 部电话, 分别在交叉口 1, 2, 5, 7.

评注　严格意义上, 集合覆盖问题有以下几个重要性质: (1) 变量 $x_j, j = 1, 2, \cdots, n$, 都是二元的; (2) 约束左端项的系数是 0 或者 1; (3) 每一个约束右端项的形式都是 $(\geqslant 1)$; (4) 目标函数是最小化 $c_1x_1 + c_2x_2 + \cdots + c_nx_n$, 其中 $c_j > 0, j = 1, 2, \cdots, n$. 在上面给出的例题中, 对所有的 j, $c_j = 1$. 如果 c_j 表示在位置 j 安装电话的费用, 那么这些系数就是这些费用值而不再是 1. 集合覆盖问题还有许多其他变形, 其中一些会在习题 8.1B 中介绍.

AMPL 程序

图 8.2 给出了求解集合覆盖问题的一般 AMPL 模型 (文件 amplEx8.2-2.txt). 一旦理解了指标集的用法 (参看附录 A.4), 就可以很容易理解这个问题的输入形式了. 模型定义了`street`作为街道的集合, 从`A`到`K`. 然后, 指标集 `corner{street}` 把街道交叉口定义为街道的函数. 有了这两个集合, 很容易建立起模型了. 图中程序的数据是关于例 8.2-1 的. 对于其他的模型, 只需要改变输入的数据.

```
#--------------Example 8.1-2--------------------
param n; #maximum number of corners
set street;
set corner{street};
var x{1..n}binary;
minimize z: sum {j in 1..n} x[j];
subject to limit {i in street}:
            sum {j in corner[i]} x[j]>=1;
data;
param n:=8;
set street:=A B C D E F G H I J K;
set corner[A]:=1 2;
set corner[B]:=2 3;
set corner[C]:=4 5;
set corner[D]:=7 8;
set corner[E]:=6 7;
set corner[F]:=2 6;
set corner[G]:=1 6;
set corner[H]:=4 7;
set corner[I]:=2 4;
set corner[J]:=5 8;
set corner[K]:=3 5;
option solver cplex;
solve;
display z,x;
```

图 8.2　关于集合覆盖问题的一般 AMPL 模型 (文件 amplEx8.2-1.txt)

习题 8.1B

***1.** ABC 是一个小型的货物配送公司, 需要每天给 5 位客户发送货物. 下表给出了每一条线路上的客户:

线路	线路上可以服务的客户
1	1, 2, 3, 4
2	4, 3, 5
3	1, 2, 5
4	2, 3, 5
5	1, 4, 2
6	1, 3, 5

由于卡车运送能力的约束, 所以每一条线路都是事先指定的. 例如, 在路线 1 上, 卡车的运送容量可以且只能满足客户 1,2,3,4 的需求. 下表给出了 ABC 总部和客户之间的距离 (单位是英里).

从 i 到 j 的距离

i \ j	ABC	1	2	3	4	5
ABC	0	10	12	16	9	8
1	10	0	32	8	17	10
2	12	32	0	14	21	20
3	16	8	14	0	15	18
4	9	17	21	15	0	11
5	8	10	20	18	11	0

目标就是找一个路程最短的日常配送方案, 以满足 5 位客户的需求. 得出的解中可能有客户会在多条选中的路线上, 但在配送执行中只选择其中的一条路线. 根据这个问题, 建立整数线性规划模型, 并求出最优解.

***2.** A 大学计划组织一个处理纠纷的委员会. 主管希望这个委员会至少包含一名女性、一名男性、一名学生、一名行政人员和一名大学教学人员. 有 10 个人 (简单标记为从字母 a 到 j) 进入候选名单, 这 10 个人可以属于上面多个类别, 如下表:

类别	个人
女性	a, b, c, d, e
男性	f, g, h, i, j
学生	a, b, c, j
行政人员	e, f
教员	d, g, h, i

A 大学希望这个委员会在满足上述要求的条件下使得人数尽量少. 根据这个问题, 建立整数线性规划模型, 并求出最优解.

3. 华盛顿县包含 6 个需要紧急救护车服务的城镇. 由于某些城镇之间距离较近, 所以一个服

务站可以服务几个城镇. 但是需要有一个约束, 那就是从服务站到它所需要服务的城镇不能超过 15 分钟的车程. 下表给出了这 6 个城镇之间的车程.

i \ j	从 i 到 j 的车程 (分钟)					
	1	2	3	4	5	6
1	0	23	14	18	10	32
2	23	0	24	13	22	11
3	14	24	0	60	19	20
4	18	13	60	0	55	17
5	10	22	19	55	0	12
6	32	11	20	17	12	0

建立整数线性规划模型并求解之. 给出服务站的最少数目以及它们的位置.

4. King Tut 要在新奥尔良的一幢博物馆举行珍宝展览, 图 8.3 给出了博物馆的地图, 不同的房间有门相通. 保安在门口可以监视两个相邻的房间. 博物馆希望每个房间都能在保安的监视中, 并且使用尽量少的保安. 根据这个问题, 建立整数线性规划模型, 并求出最优解.

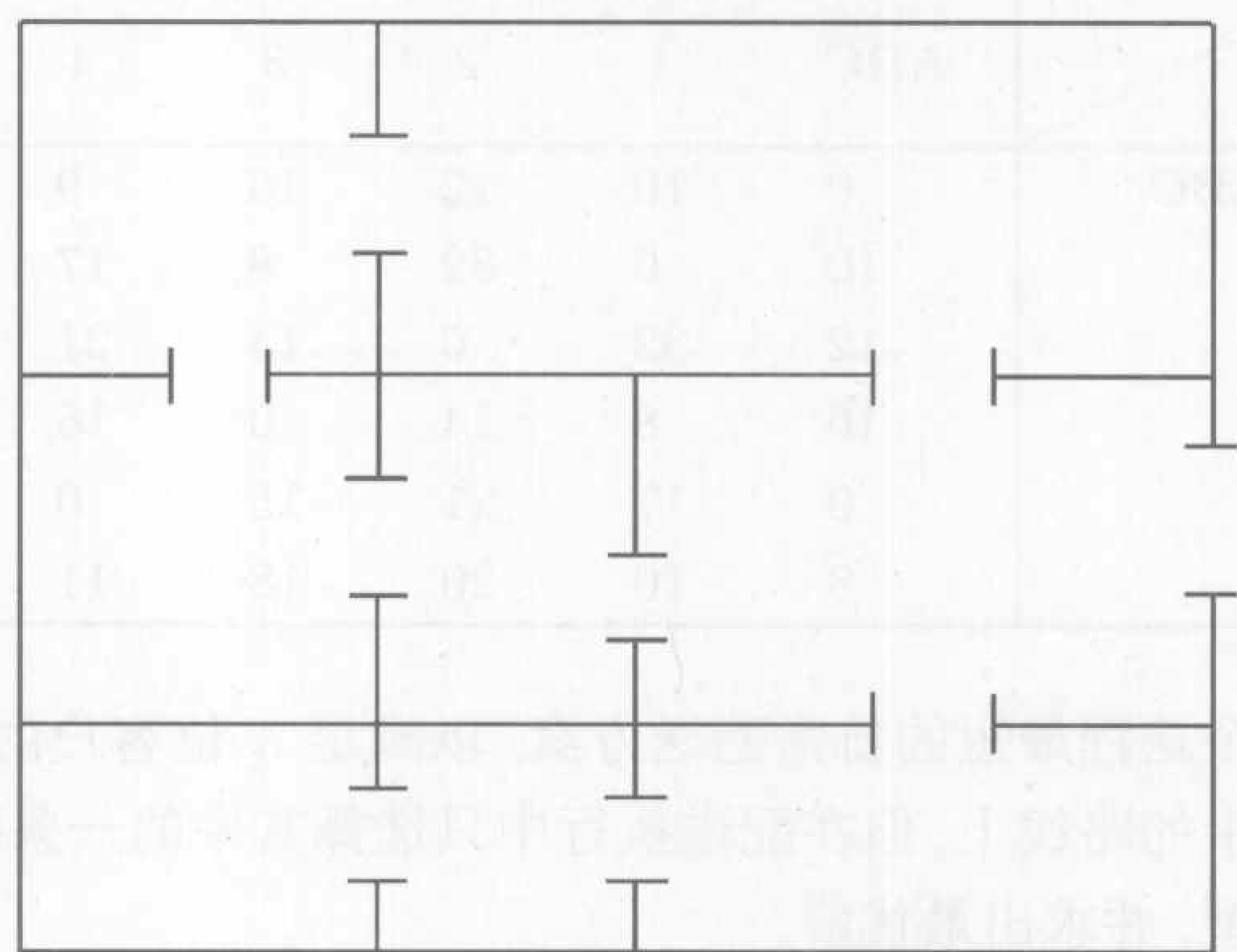


图 8.3 习题 8.1C 第 4 题中的博物馆地图

5. 为了庆祝顺利完成期末考试, Bill 打算到所在的城镇以及相邻的 6 个城镇的电影院看完所有电影. 如果他到了某个城镇, 那么他将看完所有他想看的电影再离开. 下表给出了每个城镇放映的电影以及到其他城镇的往返距离.

电影院所在城镇	放映的电影	往返距离 (英里)	每一部电影的门票
本镇	1,3	0	7.95
城镇 A	1,6,8	25	5.50
城镇 B	2,5,7	30	5.00
城镇 C	1,8,9	28	7.00
城镇 D	2,4,7	40	4.95
城镇 E	1,3,5,10	35	5.25
城镇 F	4,5,6,9	32	6.75

开车每英里的费用是 75 美分. Bill 要确定需要去的城镇, 使得可以看完每一部电影并且整个费用最小.

6. Walmark 商店正在筹划向美国西部扩展业务. 商店打算下一年开设一些新的分店, 来满足 10 个地理位置相对分散的社区. 根据以往的经验, 为了吸引顾客, 商店到社区的距离不能超过 25 英里. 另外社区人口也是开设分店时需要考虑的一个重要因素, 实际上, 社区人口越大, 能够吸引到的顾客也就越多. 下表给出了所有社区的人口以及相互之间的距离 (英里).

	从社区 i 到社区 j 的距离 (英里)									
i \ j	1	2	3	4	5	6	7	8	9	10	人口
1		20	40	35	17	24	50	58	33	12	10 000
2	20		23	68	40	30	20	19	70	40	15 000
3	40	23		36	70	22	45	30	21	80	28 000
4	35	68	36		70	80	24	20	40	10	30 000
5	17	40	70	70		23	70	40	13	40	40 000
6	24	30	22	80	23		12	14	50	50	30 000
7	50	20	45	24	70	12		26	40	30	20 000
8	58	19	30	20	40	14	26		20	50	15 000
9	33	70	21	40	13	50	40	20		22	60 000
10	12	40	80	10	40	50	30	50	22		12 000

考虑到距离的限制以及人口的集中两大因素, 开设尽可能少的分店. 指出分店应该开设在哪些社区.

***7.** (Guéret 等人, 2002, 12.6 节)MobileCo 公司拿出 1 500 万美元, 最多建造 7 个发射台来覆盖 15 个相邻社区中尽可能多的人口. 下表中给出了每个发射台可以覆盖的社区以及建造这个发射台的费用.

发射台	所覆盖社区	建造费用 (100 万美元)
1	1, 2	3.60
2	2, 3, 5	2.30
3	1, 7, 9, 10	4.10
4	4, 6, 8, 9	3.15
5	6, 7, 9, 11	2.80
6	5, 7, 10, 12, 14	2.65
7	12, 13, 14, 15	3.10

下表给出了各个社区的人口数目.

社区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
人口 (1 000)	4	3	10	14	6	7	9	10	13	11	6	12	7	5	16

确定出哪几个发射台需要建造.

8. (Gavernini 等人, 2004) 在现代电网中, 使用自动电子读表仪器代替了耗费大量人力的人工读表系统. 在自动化系统中, 几个客户的仪表通过无线的方式连接到一个接收器, 仪表每个月将客户消费的电量度数通过信号发送给指定的接收器, 然后接收器会将用户的消费信息传送给中心的计算机, 接着产生电费账单. 问题就是设置最少的接收器来服务一定数目的客户. 在现实生活中, 问题会包含成千上万的仪表和接收器. 然而为了研究的简便, 我们考虑有 10 个仪表和 8 个接收器的情形, 下表给出了相关信息:

接收器	1	2	3	4	5	6	7	8
仪表	1,2,3	2,3,9	5,6,7	7,9,10	3,6,8	1,4,7,9	4,5,9	1,4,8

求满足对所有仪表进行服务的最少数目的接收器.

9. 在第 8 题中, 如果每个接收器最多只能服务 3 个仪表, 再求解此问题.

8.1.3 固定费用问题

固定费用问题是处理一类同时包含两种费用形式的经济活动: 一种费用称作"固定费", 只要启动这种活动就会有的一个费用值; 另一种费用是可变费用, 正比于使用这种活动的程度. 举例来说, 在生产某种产品之前, 需要购买一台机器, 这台机器的费用是固定费, 它与生产多少产品无关; 一旦买进了机器, 那么劳动力和原材料的消耗费用就正比于生产产品的数量. 假定 F 是固定的费用, c 是变量的单位费用, x 是产品的数量, 那么总费用函数可以表示为

$$C(x)=\begin{cases}F+cx, & \text{如果 } x>0\\ 0, & \text{如果 } x\leqslant 0\end{cases}$$

由于费用函数 $C(x)$ 存在一个不连续点 $x=0$, 所以用解析的方法难以处理. 我们将通过下面一个例子说明如何引入二元变量来处理这类问题.

例 8.1-3 选择电话服务公司

在美国先后有 3 家电话服务公司找我推销其长途电话业务. MaBell 公司收取每月固定费 16 美元然后每分钟 0.25 美元; PaBell 公司每月收取固定费 25 美元, 但每分钟费用降为 0.21 美元; BabyBell 公司的每月固定费用是 18 美元, 每分钟 0.22 美元. 一般情况下, 我平均每月使用的长途电话时间是 200 分钟. 假定只有在我拨打电话之后公司才会收取固定费. 当然也可以同时使用多家电话公司. 那么我应该如何选择这 3 家电话公司, 使得每月的电话费用最少?

这个问题不使用整数线性规划也能求解. 然而, 可以用它作为一个例子来研究整数规划.

定义

$x_1=$ 每个月使用 MaBell 公司的长途电话时间 (分钟)

$x_2=$ 每个月使用 PaBell 公司的长途电话时间 (分钟)

$x_3=$ 每个月使用 BabyBell 公司的长途电话时间 (分钟)

$$y_i = \begin{cases} 1, & \text{如果 } x_i > 0 \\ 0, & \text{如果 } x_i = 0 \end{cases}, \quad i = 1, 2, 3$$

可以使用下面的约束来保证当 x_j 取正数时 y_j 等于 1.

$$x_j \leqslant M y_j, \ j = 1, 2, 3$$

其中 M 取足够大的一个数, 使得不会限制变量 x_j 的取值. 由于我每月使用的长途电话时间大约为 200 分钟, 所以对所有的 j, 有 $x_j \leqslant 200$, 因此取 $M = 200$ 即可.

完整的模型是

$$\begin{aligned}
\min\ & z = 0.25x_1 + 0.21x_2 + 0.22x_3 + 16y_1 + 25y_2 + 18y_3 \\
\text{s.t.}\ & x_1 + x_2 + x_3 = 200 \\
& x_1 \leqslant 200y_1 \\
& x_2 \leqslant 200y_2 \\
& x_3 \leqslant 200y_3 \\
& x_1, x_2, x_3 \geqslant 0 \\
& y_1, y_2, y_3 = (0, 1)
\end{aligned}$$

从上面的式子可以看出, 只有当 $y_j = 1$ 的时候, 也就是当 $x_j > 0$ 的时候, 第 j 个电话公司的固定费才会在目标函数中起作用 (根据模型的最后 3 个约束). 如果在最优解中 $x_j = 0$, 那么由于目标函数 z 是最小化, 又因为相应 y_j 的系数是严格的正数, 以及 $y_j \geqslant 0$, 所以一定有 $y_j = 0$ 才能达到最小.

最优解是 $x_3 = 200, y_3 = 1$, 其他的变量都取值为零, 可以看出选择 BabyBell 公司承担我的长途电话费用最为合适. 再来看最优解给出的信息, 由于 $x_3 > 0 (= 200)$ 暗示了 $y_3 = 1$, 所以 $y_3 = 1$ 是多余的. 实际上, 引入 y_1, y_2, y_3 的目的是为了计算每月的固定费. 然而就是这 3 个二元变量将一个非线性的模型转化成了一个易处理的模型. 这种使用整数 (二元) 变量来对模型进行转化的方法还可以广泛地应用到其他一些连续的问题中.

习题 8.1C

1. 皮革制品厂按合约生产一批裤子、汗衫和夹克. 每种产品都需要购买一台特定的机器来进行加工. 下表中给出了相关的数据, 包括所需要的原材料 (皮革)、生产所需的工时、生产费用以及大约的收益. 当前可以提供的皮革大约是 3 000 平方英尺, 可用的工时限制在 2 500 小时.

	裤子	汗衫	夹克
每件需要的皮革 (平方英尺)	5	3	8
每件需要的工时 (小时)	4	3	5
每件需要的生产费用 (美元)	30	20	80
每种产品需要购买设备的费用 (美元)	100	80	150
每件的价格 (美元)	60	40	120
每种产品的最少需求数目	100	150	200

求这个皮革制品厂最优的生产各种产品的数目.

***2.** Jobco 计划在 3 台机器上生产至少 2 000 个小零件. 任何一台机器上至少生产 500 个. 下表给出了这个问题的相关数据.

机器	固定费用	每件产品的生产费用	生产能力 (件)
1	300	2	600
2	100	10	800
3	200	5	1 200

根据这个问题, 建立整数线性规划模型, 并求出最优解.

***3.** Oilco 公司准备对 2 个位置进行勘测, 来检查 4 个目标 (可能是油田等). 下表提供了相关勘测的费用, 首先是在 2 个位置提前做准备的花费, 然后就是在位置 i 勘测目标 j 的费用 $(i=1,2;\ j=1,2,3,4)$. (单位: 100 万美元)

位置	每个目标的勘测费用				预先准备的费用
	1	2	3	4	
1	2	1	8	5	5
2	4	6	3	1	6

根据这个问题, 建立整数线性规划模型, 并求出最优解.

4. 考虑在 3 个工业区建制造厂, 制造厂需要给 3 个客户提供服务. 下表给出了制造厂的供应、客户的需求, 以及从制造厂到客户的单位运输费用.

制作厂 \ 客户	单位运输费用 (美元) 1	2	3	供应量
1	10	15	12	1 800
2	17	14	20	1 400
3	15	10	11	1 300
需求量	1 200	1 700	1 600	

除了运输的费用外, 修建制造厂 1, 2, 3 还需要一个固定费用, 分别是 12 000 美元、11 000 美元和 12 000 美元. 根据这个问题, 建立整数线性规划模型, 并求出最优解.

5. 对于第 4 题, 假定客户 2 和 3 的需求分别改变为 800, 再求解这个问题.

6. (Liberatore and Miller, 1985) 在随后的 6 个月, 一套加工设备使用两条生产线加工生产 3 种产品. 要求不可以违背订单, 但是为了满足下一个月的需求, 可以增加库存量. 下表给出了需求量、生产量以及 3 种产品库存量的相关数据.

产品	每个阶段的需求量						每月的单位存储费用 (美元)	初始的库存量
	1	2	3	4	5	6		
1	50	30	40	60	20	45	0.50	55
2	40	60	50	30	30	55	0.35	75
3	30	40	20	70	40	30	0.45	60

从生产一种产品的生产线转到生产另一种产品的生产线需要一个固定的费用. 下表给出了每条生产线的转换费用、生产效率以及每种产品的单位生产成本:

	生产线转换费用 (美元)		
	产品 1	产品 2	产品 3
生产线 1	200	180	300
生产线 2	250	200	174

	生产效率 (每月件数)			单位生产费用 (美元)		
	产品 1	产品 2	产品 3	产品 1	产品 2	产品 3
生产线 1	40	60	80	10	8	15
生产线 2	90	70	60	12	6	10

根据这个问题, 建立整数线性规划模型, 并求出最优的生产计划.

7. (Jarvis 等人, 1978)7 个城市被候选为建设 4 个污水处理厂的城市. 下表给出了相关信息, 表中没有给出的连接说明这条线路上不能铺设管道.

从 \ 到	1	2	3	4	5	6	7
	两个城市之间建造流量为 1 000 加仑/小时管道的费用 (美元)						
1		100		200		50	
2				120		150	
3	400				120		90
4			120		120		
5		200				100	200
6			110	180			70
7	200			150			
建造工厂的费用 (100 万美元)	1.00	1.20	2.00	1.60	1.80	0.90	1.40
人口 (1 000 人)	50	100	45	90	75	60	30

管道的容量 (加仑/小时) 是产生污水数量的函数, 也就是人口数目的函数. 大约每小时每 1 000 居民向下水道排放的污水是 500 加仑. 污水处理厂最大处理能力是 10 万加仑/小时. 试确定建设污水处理厂的最佳位置及其处理能力.

8. (Brown 等人, 1987) 某公司使用 4 辆特殊的油罐卡车为客户运送 4 种不同的汽油. 每辆油罐车有 5 节车厢, 容量分别是 500, 750, 1 200, 1 500, 1 750 加仑. 4 种汽油每天的需求量大约分别是 10 000, 15 000, 12 000, 8 000 加仑. 4 辆油罐车运送的不能满足客户需求的部分必须转包合同, 那么 4 种汽油的额外转包费用分别是 5, 12, 8, 10 美分/加仑. 求最小化转包合同产生的额外费用时的最优油罐车日常运输计划.

9. 某家庭每月需要的长途电话大约是 3 000 分钟, 现在有 3 家公司可以提供电话服务: A, B, C. 公司 A 收取每月的固定费 10 美元, 然后前 1 000 分钟 5 美分/分钟, 以后超出 1 000 分钟的部分 4 美分/分钟; 公司 B 每月收取固定费 20 美元, 长途费用统一为 4 美分/分

钟；公司 C 的每月固定费为 25 美元, 然后前 1 000 分钟 5 美分/分钟, 以后超出 1 000 分钟的部分 3.5 美分/分钟. 那么这户家庭应该选择哪家公司使得电话费用最少.

***10.** (Barnett, 1987)Yataha 教授计划有 6 次波士顿与华盛顿之间的往返旅行. 可以选择的 3 条定期航线是 Eastern, US Air 和 Continental, 并且购买单程机票也不会额外收钱. 为了吸引旅客, 每条航空线路都对经常乘坐客机的乘客有奖励里程. Eastern 给予每张机票(单程)1 000 英里奖励; 另外如果每月的机票数量达到 2 张, 又会额外赠送 5 000 英里; 如果数量超过 5 张将再赠送 5 000 英里. Air 给予每次旅行奖励 1 500 英里, 每 6 张机票额外赠送 10 000 英里. Continental 给予每次旅行奖励 1 800 英里, 每 5 张机票额外赠送 7 000 英里. 那么 Yataha 教授应该如何在 3 条航线上分配这 12 张单程机票, 才能使得自己得到的奖励里程最大?

8.1.4 “或者–或者”和“如果–那么”约束

在固定费用问题 (8.1.3 节) 中, 我们引入二元变量来处理不连续的目标费用函数. 在本节中, 我们将仍然利用二元变量来处理模型中约束不满足同时性 (“或者–或者”) 或者依赖性 (“如果–那么”) 的情形. 采用的变换不会改变约束的“或者”或“依赖”关系, 我们只是利用数学上的技巧方法将这样的约束转化成“并且”的关系.

例 8.1-4 (工序模型)

Jobco 公司需要在一台机器上处理 3 项工作. 下表给出了每项工作的处理时间以及应交工日期. 假定第一项工作开始处理时的日期定为 0, 应交工日期从 0 算起.

工作	处理时间 (天)	应交工日期 (天)	延期处罚 (美元/天)
1	5	25	19
2	20	22	12
3	15	35	34

这个问题的目标是求使得延期处罚最少的工序处理方案.

定义 $x_j =$ 工作 j 的开始加工日期 (按天计算, 从 0 开始计算)

这个问题有两个约束条件: 互不干扰约束 (保证两项工作不能同时处理), 应交工日期约束. 首先考虑互不干扰约束.

假设工作 i 和工作 j 的处理时间分别为 p_i 和 p_j, 为了保证这两项工作不同时处理, 那么必须满足 $x_i \geqslant x_j + p_j$ 或者 $x_j \geqslant x_i + p_i$, 取决于工作 j 是在工作 i 之前处理还是在其后处理. 因为所有的数学规划都是仅仅处理并的约束, 所以我们通过引入下面的二元变量来将“或者–或者”约束转化为“并的”约束:

$$y_{ij} = \begin{cases} 1, & \text{若工作 } i \text{ 在工作 } j \text{ 之前加工} \\ 0, & \text{若工作 } j \text{ 在工作 } i \text{ 之前加工} \end{cases}$$

对于足够大的数 M, **或者–或者约束** (either-or constraint) 可以用下面两个并的约束代替 $My_{ij} + (x_i - x_j) \geqslant p_j$ 和 $M(1 - y_{ij}) + (x_j - x_i) \geqslant p_i$

这样的转化可以保证在任何时候这两个约束中只有一个是起作用的. 如果 $y_{ij}=0$, 那么第 1 个约束是起作用的, 而第 2 个约束是多余的 (因为左端项包含充分大的数 M, 所以一定会大于 p_i); 如果 $y_{ij}=1$, 那么第 1 个约束是多余的, 第 2 个约束是起作用的.

下面再考虑应交工日期约束. 给定工件 j 的应交工日期是 d_j, 令 s_j 是一个无限制的变量. 那么相关的约束是

$$x_j+p_j+s_j=d_j$$

如果 $s_j\geqslant 0$, 那么应交工日期的约束可以满足; 如果 $s_j\leqslant 0$ 则会有一个延期的处罚. 引用下面的代换:

$$s_j=s_j^- - s_j^+,\ \ s_j^-,s_j^+\geqslant 0$$

那么约束变为

$$x_j+s_j^- - s_j^+=d_j-p_j$$

延期处罚的费用与 s_j^+ 成正比.

这个问题的模型可以写成

$$\begin{aligned}
\min\ & z=19s_1^+ +12s_2^+ +34s_3^+\\
\text{s.t.}\ & x_1-x_2+My_{12}\geqslant 20\\
& -x_1+x_2-My_{12}\geqslant 5-M\\
& x_1-x_3+My_{13}\geqslant 15\\
& -x_1+x_3-My_{13}\geqslant 5-M\\
& x_2-x_3+My_{23}\geqslant 15\\
& -x_2+x_3-My_{23}\geqslant 20-M\\
& x_1+s_1^- - s_1^+=25-5\\
& x_2+s_2^- - s_2^+=22-20\\
& x_3+s_3^- - s_3^+=35-15\\
& x_1,x_2,x_3,s_1^-,s_1^+,s_2^-,s_2^+,s_3^-,s_3^+\geqslant 0\\
& y_{12},y_{13},y_{23}=(0,1)
\end{aligned}$$

引入整数变量 y_{12}, y_{13}, y_{23} 是为了将“或者–或者”约束转化成“并的”约束. 最后构造的模型是一个混合整数线性规划模型.

为了求解这个模型, 我们取 $M=100$, 这个值比 3 项工作处理时间的总和还要大.

最优解是 $x_1=20$, $x_2=0$ 和 $x_3=25$. 这说明工作 2 在时刻 0 开始处理, 工作 1 在第 20 天开始处理, 工作 3 在第 25 天开始处理, 最优的处理顺序是 $2\to 1\to 3$. 同时这个解还告诉我们工作 2 在第 $0+20=20$ 天完成, 工作 1 在第 $20+5$ 天完成, 工作 3 在第 $25+15=40$ 天完成, 所以工作 3 延误了 $40-35=5$ 天, 处罚费用为 $5\times 34=170$ 美元.

AMPL 程序

文件 amplEx8.1-4.txt 为例 8.1-4 提供了 AMPL 模型. 由于它是对上面提出的一般数学规划模型的直接转化, 所以这个模型很容易理解. 通过改变数据的输入, 这个模型可以处理任意多项工作的问题. 注意, 实际上这个模型可以看作是下面几个原始数据量的函数：处理时间p、截止时间d和延期处罚 perDaypenalty.

```
#------------------Example 8.1-4-------------------
param n;
set I={1..n};
set J={1..n};    #I is the same as J
param p{I};
param d{I};
param perDayPenalty{I};
param M=1000;
var x{J}>=0;           #continuous
var y{I,J} binary;   #0-1
var sMinus{J}>=0;    # s=sMinus-sPlus
var sPlus{J}>=0;
minimize penalty: sum {j in J}
          perDayPenalty[j]*sPlus[j];
subject to
eitherOr1{i in I,j in J:i<>j}:
       M*y[i,j]+x[i]-x[j]>=p[j];
eitherOr2{i in I,j in J:i<>j}:
       M*(1-y[i,j])+x[j]-x[i]>=p[i];
dueDate{j in J}:x[j]+sMinus[j]-sPlus[j]=d[j]-p[j];
data;
param n:=3;
param p:= 1 5   2 20  3 15;
param d:= 1 25  2 22  3 35;
param perDayPenalty := 1 19  2 12 3 34;
option solver cplex; solve;
display penalty,x;
```

图 8.4　工序问题的 AMPL 模型 (文件 amplEx8.1-4.txt)

例 8.1-5 (再次考虑工序模型)

在例 8.1-4 中, 假定我们有下面的额外约束：如果工作 i 在工作 j 之前, 那么工作 k 必须在工作 m 之前. 从数学角度看, 这个**如果–那么约束** (if-then condition) 可以写成

$$\text{如果 } x_i + p_i \leqslant x_j, \text{ 那么 } x_k + p_k \leqslant x_m$$

给定足够小的数 $\varepsilon > 0$ 和足够大的数 M, 上面这个约束等价于下面 2 个并的约束：

$$x_j - (x_i + p_i) \leqslant M(1 - w) - \varepsilon$$

$$(x_k + p_k) - x_m \leqslant Mw$$

$$w = (0, 1)$$

如果 $x_i + p_i \leqslant x_j$, 那么 $x_j - (x_i + p_i) \geqslant 0$, 这要求 $w = 0$, 因此第 2 个约束就变成了 $x_k + p_k \leqslant x_m$, 这正是我们需要的. 否则, w 取值为 1 或者 0, 此时第 2 个式子是否满足, 取决于模型中其他的约束.

习题 8.1D

***1.** 一个游戏板上由 9 个大小相同的正方形组成, 要求将数 1 ~ 9 分别填入这些正方形, 每个正方形填一个数而且互不相同, 使得每一行、每一列以及对角线上的数的和都是 15. 利用整数线性规划模型来确定应该如何将这 9 个数填写到所有的正方形中.

2. 一台机器可以用来生产 2 种可以转换的产品. 这台机器每天最多可以生产 20 个单位的产品 1 和 10 个单位的产品 2, 或者可以调整为每天最多生产 12 个单位的产品 1 和 25 个单位的产品 2. 市场的分析数据显示每天对两种产品的最大需求量一共是 35 个单位. 已知 2 种产品的单位利润分别是 10 美元和 12 美元, 那么应该选择哪种生产模式呢? 根据这个问题, 建立整数线性规划模型, 并求出最优解. (注: 这是一个二维的问题, 可以通过检查图上可行解空间的方法求解, 但是这种方法不适用于求解 n 维的情况.)

***3.** Gapco 公司生产加工 3 种产品, 下表给出了它们需要的加工工时以及所需要的原料.

产品	每天需要的工时 (小时/单位)	每天需要的原料 (磅/单位)
1	3	4
2	4	3
3	5	6

这 3 种产品每单位的利润分别是 25 美元、30 美元和 22 美元. 有 2 个位置可以提供给 Gapco 来建造工厂, 这 2 个位置可以提供不同的生产工时和原材料数量, 见下表:

位置	每天可用工时 (小时)	每天可用原料 (磅)
1	100	100
2	90	120

根据这个问题, 建立整数线性规划模型, 并求出最优的建厂位置.

4. Jobco 车间需要在一台机器上处理 10 项主要的工作. 下表给出了每项工作的处理时间以及应交工日期. 所有的时间单位为天, 日期从 0 开始.

工作	处理时间	应交工日期
1	10	20
2	3	98
3	13	100
4	15	34
5	9	50
6	22	44
7	17	32
8	30	60
9	12	80
10	16	150

如果工作 4 在工作 3 之前处理, 那么工作 9 必须在工作 7 之前处理. 目标是在最短的时间内将这 10 项工作处理完毕. 根据这个问题, 建立整数线性规划模型, 并利用文件 amplEx8.1-4.txt 中的 AMPL 模型进行求解.

5. 在第 4 题中, 假定只有在工作 3 完成以后才能够处理工作 4, 并且机器的设置要求工作 7 和工作 8 必须相继加工 (也就是, 要么工作 7 在工作 8 之后马上加工, 要么工作 8 在工作 7 之后马上加工). Jobco 公司的目标是使得处理所有的 10 项工作的总延误时间最少. 根据这个问题建立数学模型, 并求出最优解.

6. Jaco 公司拥有一家工厂, 生产 3 种产品. 下表给出了这 3 种产品需要的工时和原材料.

产品	每天需要的工时 (小时/单位)	每天需要的原料 (磅/单位)
1	3	4
2	4	3
3	5	6
每天可用量	100	100

这 3 种产品的单位利润分别是 25 美元、30 美元和 45 美元. 如果生产产品 3, 那么每天至少要生产 5 件. 根据这个问题, 建立一个混合整数线性规划模型, 并求出最优的生产方案.

7. UPak 公司是 LTL 运输公司的一家子公司. 客户将货物送到 UPak, 然后装入拖车运送到客户需要的目的地, 拖车可以出租的空间是 36 英尺. 客户根据他们货物所占英尺数付给 UPak 费用. 客户的货物不能分开装车, 也就是说每一个客户的货物必须装在一辆拖车上. 在拖车上安装了可移动的隔板用来分开不同客户的货物. UPak 对每英尺空间收取的费用取决于货物的目的地: 距离越远, 费用也就越高. 下表中给出了 UPak 公司需要运输的货物单据.

单据	1	2	3	4	5	6	7	8	9	10
货物尺寸 (英尺)	5	11	22	15	7	9	18	14	10	12
单位运费 (美元)	120	93	70	85	125	104	98	130	140	65

现在有两辆拖车等待装货, 确定如何选择货物, 使得总的收入最大. (*提示*: 引入二元变量 x_{ij} 表示第 i 个货物装到第 j 个拖车上. 关键是定义在拖车 j 上装货物 i 的变量 x_{ij}, 需要有一个*如果–那么*的约束来避免货物的分装.)

8. 分别对图 8.5 中的 3 个图形的阴影可行解空间用一组并的约束表示. 对于图 (a) 给出的可行解空间求使得 $z = 2x_1 + 3x_2$ 最大化的解.

9. 假定要求下面 m 个约束中的*任何k* 个一定起作用:

$$g_i(x_1, x_2, \cdots, x_n) \leqslant b_i,\ i = 1, 2, \cdots, m$$

请给出如何表示这样的约束.

10. 在下面的约束中, 右端项 $(b_2, b_2, \cdots,$ 或 $b_m)$ 代表取 $b_1, b_2, \cdots, b_m$ 中的一个.

$$g(x_1, x_2, \cdots, x_n) \leqslant (b_1, b_2, \cdots, \text{或 } b_m)$$

请给出如何表示这样的约束.

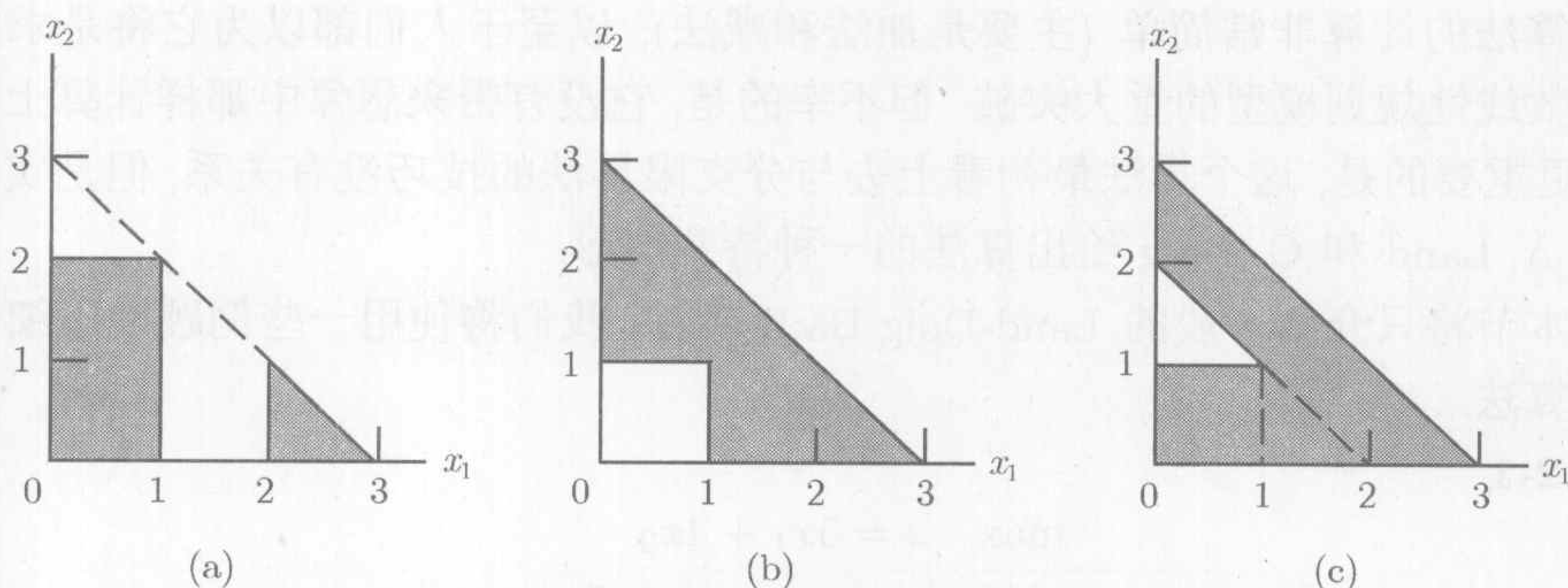


图 8.5 习题 8.1D 中第 9 题的可行解空间

8.2 整数规划算法

求解整数线性规划的算法大都是建立在求解大规模线性规划问题的算法基础之上的. 这种算法的策略包含 3 个步骤.

第 1 步 首先对整数线性规划的可行解空间进行松弛: 取消对整数变量的整数限制, 用连续的区间 $0 \leqslant y \leqslant 1$ 来代替二元变量 y. 松弛之后的结果是一个规则的线性规划模型.

第 2 步 求解这个线性规划模型, 并得到连续最优解.

第 3 步 根据得到的连续最优解, 通过增加一些特殊的约束来逐步改变线性规划的可行解空间, 使得最终得到一个满足整数要求的最优极点.

现在有两种常用的方法来产生第 3 步中的特殊约束:

(1) 分支限界法; (2) 割平面法.

虽然两种方法都不能够保证对所有的整数规划有效, 但经验表明分支限界法远远好于割平面法. 我们会在接下来的章节中进一步加以讨论.

8.2.1 分支限界 (B&B) 算法

1960 年 A. Land 和 G. Doig 设计了第一个用于求解一般的混合整数线性规划和纯整数线性规划问题的分支限界算法①. 接下来, 在 1965 年, E. Balas 又得到解决纯二元 (0 或者 1) 变量② 组成的整数线性规划问题的**附加算法** (additive algorithm).

① TORA 的整数规划模块包含一个交互式的生成 B&B 树的工具. 使用这个工具的时候, 选择整数规划模块输出屏幕上的 User-guided B&B 按钮, 然后屏幕显示结果提供了所有生成 B&B 树的信息.

② 一般的整数线性规划模型的变量都可以转化成二元 $(0-1)$ 变量, 给定整数型变量 x, 有有限的上界 u(也就是 $0 \leqslant x \leqslant u$), 那么

$$x = 2^0 y_0 + 2^1 y_1 + 2^2 y_2 + \cdots + 2^k y_k$$

变量 $y_0, y_1, \cdots, y_k$ 都是二元变量, 下标 k 是满足 $2^{k+1} - 1 \geqslant u$ 的最小整数.

附加算法的计算非常简单 (主要是加法和减法), 以至于人们都以为它将是求解一般整数线性规划模型的重大突破. 但不幸的是, 它没有带来想象中那样计算上的优势. 更重要的是, 这个算法最初看上去与分支限界法的技巧没有关系, 但是实际上它是 A. Land 和 G. Doig 给出算法的一种特殊情形.

本节将只介绍一般的 Land-Doig B&B 算法, 我们将使用一些例题来仔细讲解这个算法.

例 8.2-1

$$
\begin{aligned}
\max \quad & z = 5x_1 + 4x_2 \\
\text{s.t.} \quad & x_1 + x_2 \leqslant 5 \\
& 10x_1 + 6x_2 \leqslant 45 \\
& x_1, x_2 \text{均是非负整数}
\end{aligned}
$$

图 8.6 中的格点 (黑点) 就是这个整数线性规划问题的可行解空间. 取消整数线性规划问题中的整数约束, 就得到了相应的第 1 个连续线性规划 LP1 问题. 可行解空间是图中的阴影部分, 求解这个 LP1 问题, 得到 $x_1 = 3.75$, $x_2 = 1.25$, $z = 23.75$.

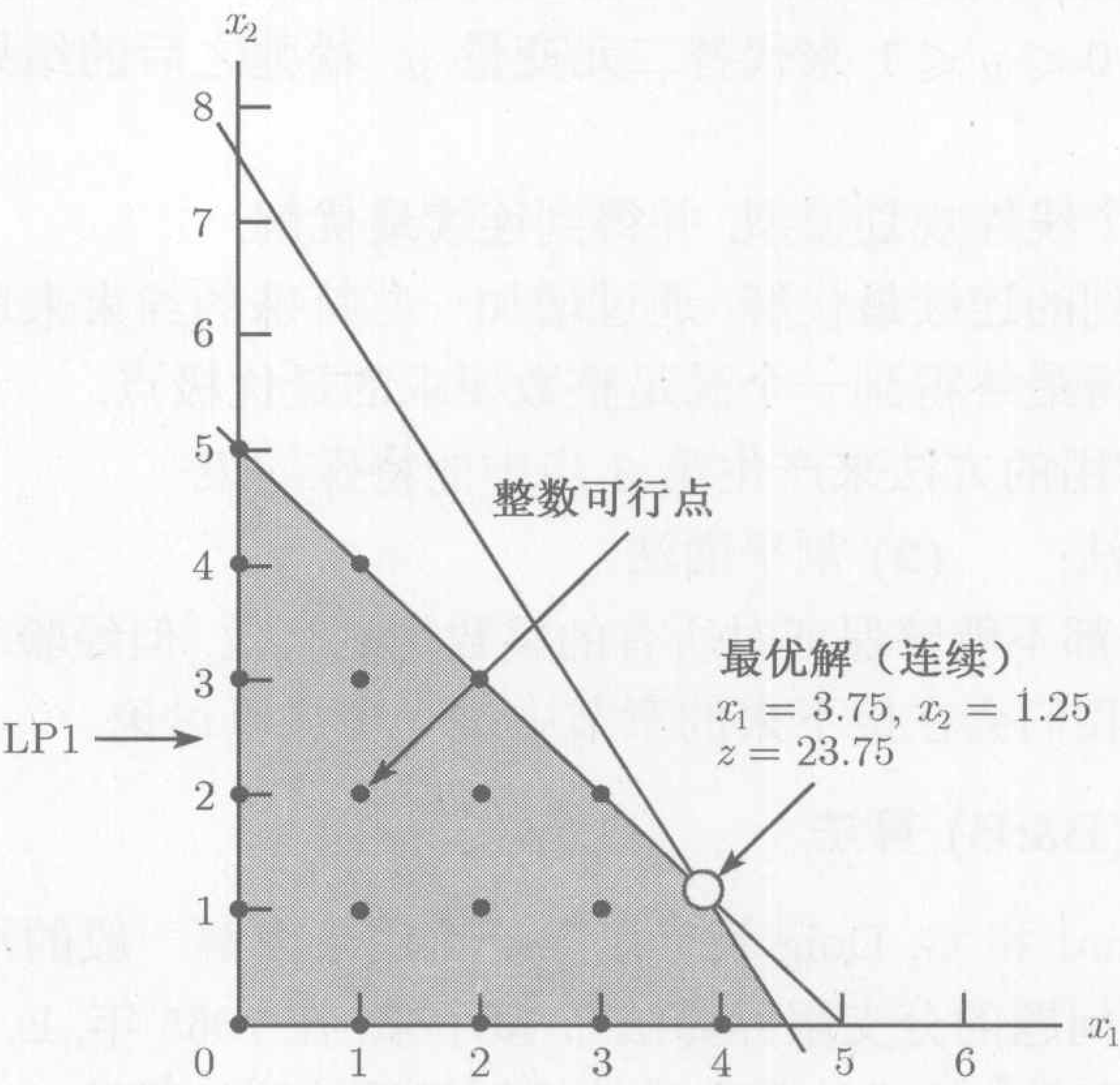


图 8.6　例 8.2-1 中整数线性规划 (格子点) 和 LP1(阴影部分) 的可行解空间

由于得到的 LP1 问题的最优解不满足整数要求, 所以使用 B&B 算法来改变可行解空间使得最终得到整数线性规划的最优解. 首先, 我们选择 LP1 最优解中任何一个不满足整数要求的变量, 例如选择 $x_1(= 3.75)$, 由于在 LP1 的可行解空间中, 区间 $3 < x_1 < 4$ 中不含有 x_1 的整数值, 所以可以不考虑变量 x_1 的这一段区域. 可以等价地用两个新的线性规划来代替原来的 LP1:

$$\text{LP2 可行解空间} = \text{LP1可行解空间} + (x_1 \leqslant 3)$$
$$\text{LP3 可行解空间} = \text{LP1可行解空间} + (x_1 \geqslant 4)$$

图 8.7 描述的 LP2 和 LP3 的可行解空间, 这两个区域中的整数点刚好与原始的整数线性规划问题的解区域相同, 也就是说, 从求整数解的角度出发, 求解 LP2 和 LP3 与求解 LP1 可以得到相同的结果, 不会丢失任何信息.

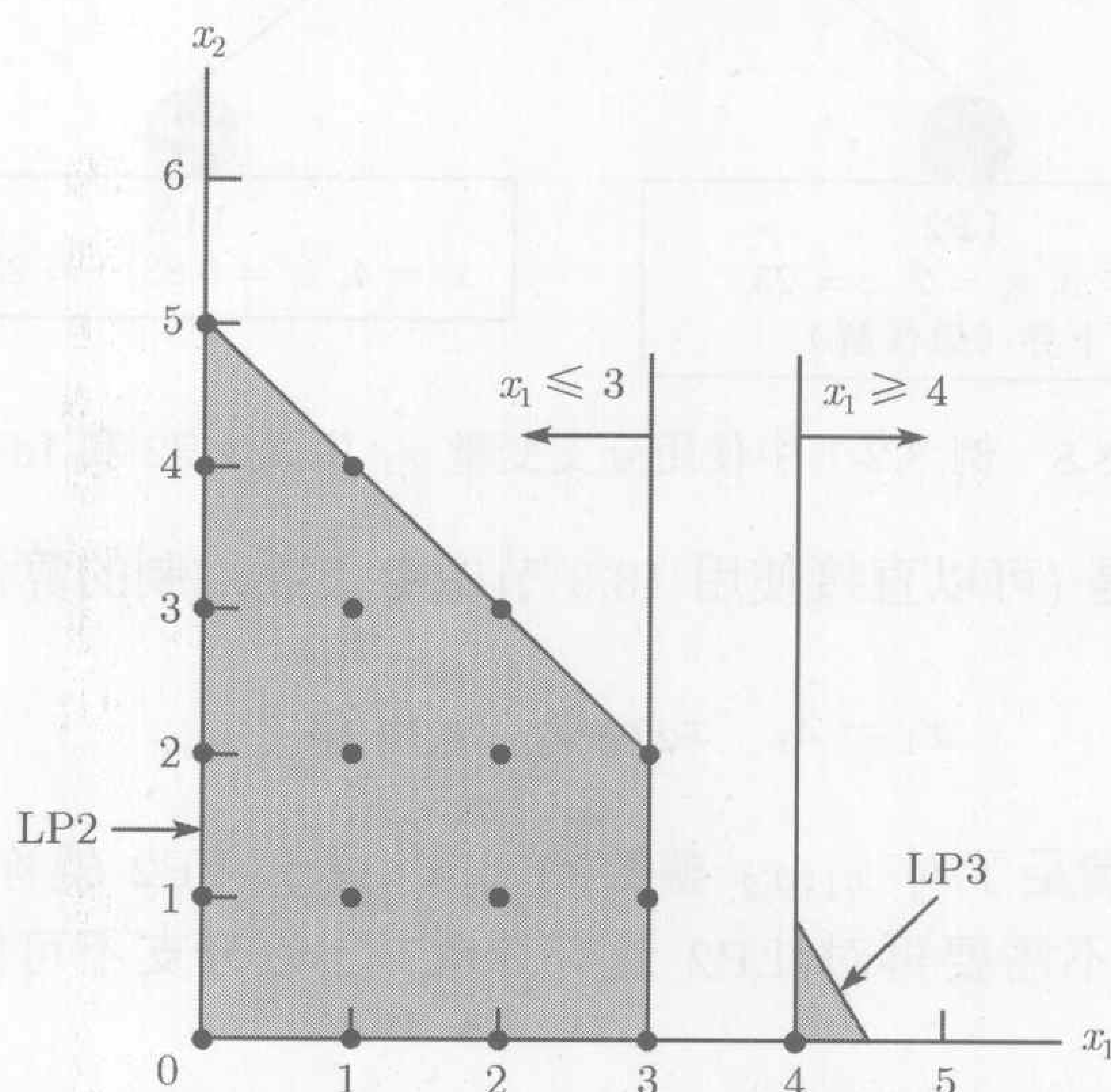


图 8.7 例 8.2-1 中问题 LP2 和 LP3 的可行解空间

由此我们可以考虑通过增加合适的约束将可行解空间内不包含整数点的区域去除 (就像从 LP1 的可行解空间中去除 $3 < x < 4$ 一样), 最终通过求解一个线性规划问题, 使得得到的解恰好满足整数要求. 这样, 我们求解整数线性规划的过程就变成了求解一系列连续的线性规划问题.

由于新的约束 $x_1 \leqslant 3$ 和 $x_1 \geqslant 4$ 是互不相交的, 所以图 8.8 给出的第 2 个节点和第 3 个节点 (LP2 和 LP3) 是可行解空间不相关的线性规划问题, 因此这种二分的性质给我们提供了采用 B&B 算法中**分支** (branching) 的思想. 在这种情况下, 变量 x_1 称为**分支变量** (branching variable).

整数线性规划的最优解或者落在 LP2 上, 或者落在 LP3 上, 因此两个子问题都需要考虑. 首先求解 LP2(引入了约束 $x_1 \leqslant 3$):

$$\begin{aligned} \max \quad & z = 5x_1 + 4x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leqslant 5 \\ & 10x_1 + 6x_2 \leqslant 45 \\ & x_1 \leqslant 3 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

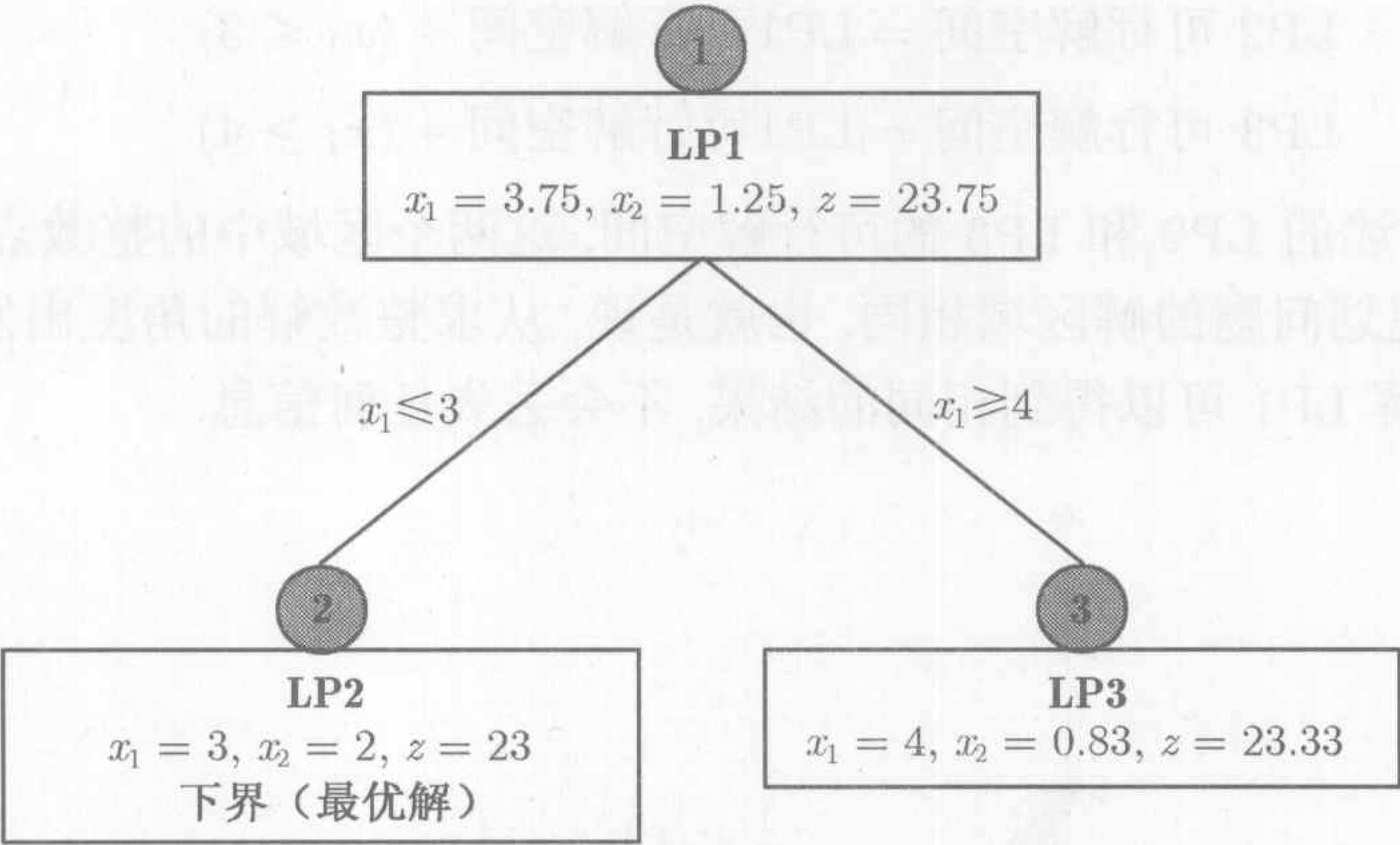


图 8.8　例 8.2-1 中使用分支变量 x_1 生成 LP2 和 LP3

LP2 的最优解是 (可以直接使用 13.3 节中有上界限制的算法来快速求解)

$$x_1 = 3, \quad x_2 = 2, \quad z = 23$$

LP2 的最优解满足了对 x_1, x_2 整数的要求, 那么 LP2 就称作**已经探明** (fathomed), 也就意味着不需要再对 LP2 进行分支, 因为分支不可能再得到*更*好的整数解.

现在还不能说 LP2 得到的整数解就是原问题的最优解, 因为求解 LP3 可能会得到更好的整数解. 我们可以得到的结论就是 $z = 23$ 是原始整数线性规划问题最优 (最大) 值的一个**下界**. 只要是不能够得到更好的目标值的子问题就可以不用再继续考虑. 如果另外一个子问题得到了更好的整数解, 那么这个下界也就随之得到提高.

已知下界 $z = 23$, 下面考虑 LP3(唯一剩下没有考虑的子问题). 因为 LP1 的最优值是 $z = 23.75$, *并且目标函数中所有的系数都是整数*, 所以求解 LP3(比 LP1 的约束更多) 不会得到更好的解使得 $z > 23$, 因此可以不用再考虑子问题 LP3, 它也*已经探明*.

由于子问题 LP2 和 LP3 都已经探明 (第一个子问题得到了最优整数解, 第二个不能够得到*更*好的整数解), 那么可以得到原始整数线性规划问题的最优解是 $x_1 = 3, x_2 = 2, z = 23$.

在整个过程中还有两个问题没有回答.

(1) 在求解 LP1 时, 是否可以选择 x_2 代替 x_1 作为*分支变量*?

(2) 当选择下一个需要求解的子问题时, 是否可以考虑先求解 LP3?

这两个问题的答案都是“是”, 但是在计算量上会有很大的差异, 图 8.9 充分说明了这一点. 假定首先考虑 LP3(而不是按照图 8.8 中首先考虑 LP2), 可以得到最

优解为 $x_1 = 4, x_2 = 0.83, z = 23.33$(已检验！). 因为 $x_2(= 0.83)$ 不是整数, 所以可以利用 $x_2 \leqslant 0$ 和 $x_2 \geqslant 1$ 将 LP3 进一步分成两个子问题 LP4 和 LP5, 也就是

$$\begin{aligned}\text{LP4 可行解空间} &= \text{LP3 可行解空间} + (x_2 \leqslant 0)\\ &= \text{LP1 可行解空间} + (x_1 \geqslant 4) + (x_2 \leqslant 0)\\ \text{LP5 可行解空间} &= \text{LP3 可行解空间} + (x_2 \geqslant 1)\\ &= \text{LP3 可行解空间} + (x_1 \geqslant 4) + (x_2 \geqslant 1)\end{aligned}$$

现在我们有 3 个子问题需要求解：LP2, LP4, LP5. 假定首先任意选取其中一个, 比如 LP5, 求解发现它没有解, 因此它已经探明. 接下来求解 LP4, 最优解是 $x_1 = 4.5, x_2 = 0, z = 22.5$, 因为 x_1 的值不是整数, 所以又得到两个分支 $x_1 \leqslant 4$ 和 $x_1 \geqslant 5$, 从 LP4 可以得到下面的两个子问题 LP6 和 LP7.

$$\begin{aligned}\text{LP6 可行解空间} &= \text{LP1 可行解空间} + (x_1 \geqslant 4) + (x_2 \leqslant 0) + (x_1 \leqslant 4)\\ \text{LP7 可行解空间} &= \text{LP3 可行解空间} + (x_1 \geqslant 4) + (x_2 \leqslant 0) + (x_1 \geqslant 5)\end{aligned}$$

现在还有 3 个子问题没有考虑：LP2, LP6, LP7. 任意选择 LP7 求解, 这个问题没有可行解, 因此它已经探明. 接着考虑 LP6, 求解这个问题可以得到整数解 $x_1 = 4, x_2 = 0, z = 20$, 也就得到了第一个原始整数线性规划最优值的下界 $(= 20)$. 下面只剩下子问题 LP2, 求解得到一个更好的整数解 $(x_1 = 3, x_2 = 2, z = 23)$. 到现在为止, 所有的子问题都是已经探明, 并且根据最新的一个下界得到原问题的最优解 $x_1 = 3, x_2 = 2, z = 23$.

图 8.9 给出了整个求解序列 (LP1→LP3→LP5→LP4 →LP7→LP6→LP2), 显然这样一个序列是一种非常糟糕的方案, 但是实际中也是会发生的. 在图 8.8 中, 我们幸运地选中了求解 LP2 得到了一个好的下界, 从而使得我们不再需要求解 LP3, 实质上只是求解了两个线性规划问题. 但是在图 8.9 中就截然不同了, 在 B&B 算法终止之前我们求解了 7 个线性规划问题.

评注 从上面的例题中可以看出, B&B 算法有一个非常大的缺点：给定多个可供选择的子问题时, 如何确定下一个求解的子问题以及分支变量？虽然也有一些启发式的想法使得 B&B 算法可以预见哪一个分支会得到一个好的整数解 (参看 Taha, 1975, pp.154-171), 但是理论上仍然没有一个统一的结果, 这也一直以来困扰着更好地求解整数线性规划. 习题 8.2A 第 7 题将证明, 即使问题的规模比较小 (16 个二元变量和 1 个约束), 利用 B&B 算法在找到最优解之前也需要求解超过 25 000 个线性规划问题. 不幸的是, 经过了 40 余年的研究, 并且在计算能力有了很大提高的前提下, 直到今天仍然没有一套完全可靠的整数线性规划代码 (商业的和学术上的), 也就是无论人们在计算机上执行多长时间都不能找到整数线性规划的最优解. 更让人沮丧的是, 这些算法甚至在一些相当小的问题上也会受到限制.

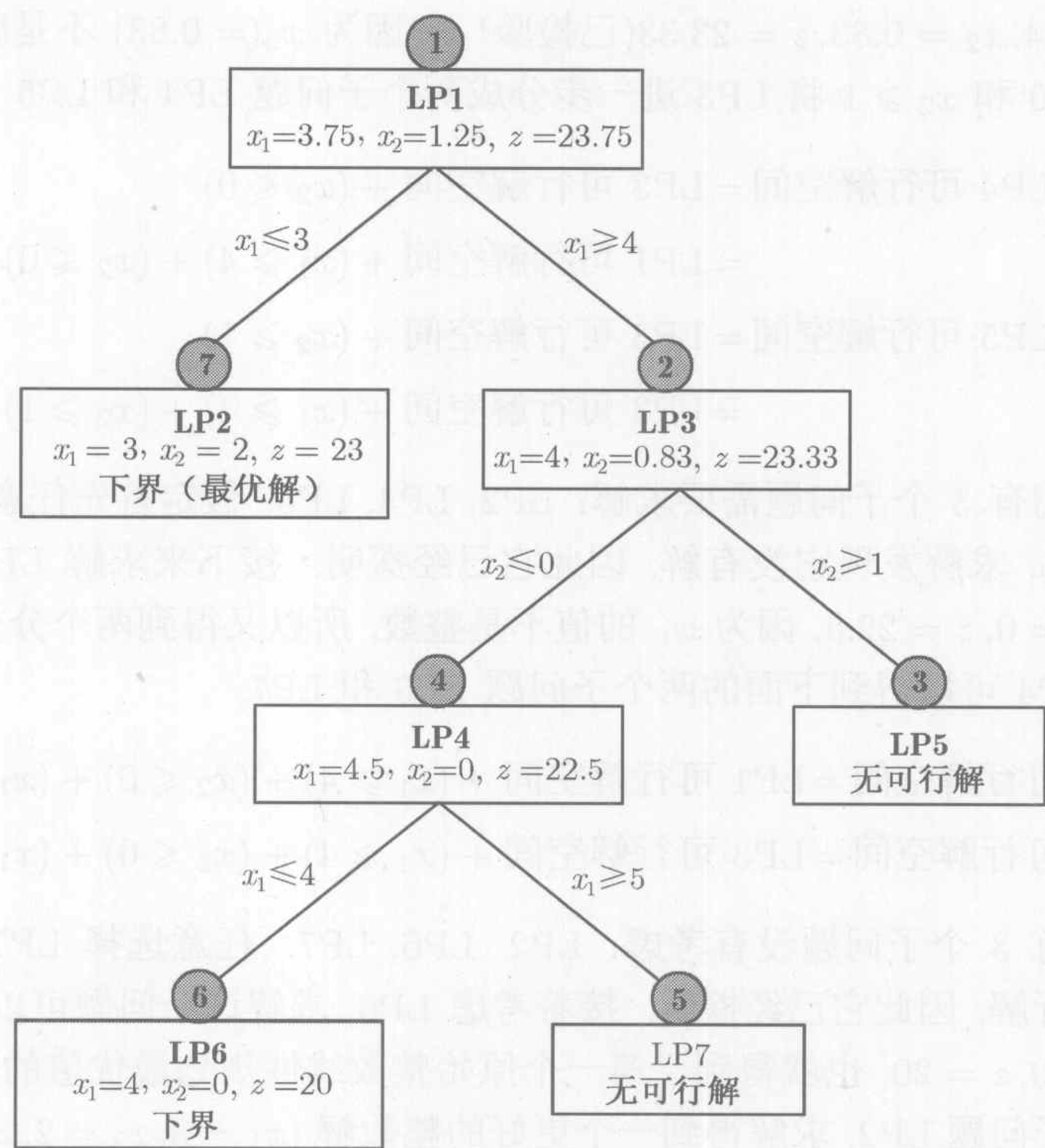


图 8.9　例 8.2-1 中的分支限界树

AMPL 模型

AMPL 模型可以用来交互式地生成 B&B 搜索树. 下表中给出了生成例 8.2-1 (图 8.9) 分支限界树所需要的命令顺序, 从连续的 LP0 模型开始. 这个模型 (文件 amplEx8.2-1.txt) 包含两个变量x1和x2, 两个约束c0和c1. 可以看出这些命令与图 8.9 给出的分支是相互对应的.

AMPL 命令	结　果
ampl: model amplEx8.2-1.txt;solve;display x1,x2;	LP1($x_1 = 3.75, x_2 = 1.25$)
ampl: c2:x1>=4;solve;display x1,x2;	LP3($x_1 = 4, x_2 = 0.83$)
ampl: c3:x2>=1;solve;display x1,x2;	LP5(无解)
ampl: drop c3;c4:x2<=0;solve;display x1,x2;	LP4($x_1 = 4.5, x_2 = 0$)
ampl: c5:x1>=5;solve;display x1,x2;	LP7(无解)
ampl: drop c5;c6:x1<=4;solve;display x1,x2;	LP6($x_1 = 4, x_2 = 0$)
ampl: drop c2;drop c4;drop c6;c7:x1<=3; solve;display x1,x2;	LP2($x_1 = 3, x_2 = 2$)

Excel 规划求解程序

在 Excel 规划求解中可以利用**规划求解参数**对话框上的 add/change/delete 选项得到想要的子问题的解.

B&B 算法总结 现在总结一下 B&B 算法. 假定一个求目标函数最大的问题, 设置初始最优目标值的下界是 $z=-\infty$. 令 $i=0$.

第 1 步 (探明/限界) 选择 LPi 为下一个需要考虑的子问题. 求解 LPi, 并按照下面的 3 种条件之一来判断这个问题是否已经探明:

(a) LPi 的最优值 z 不能改进目标函数当前的下界;

(b) LPi 得到了一个比当前目标函数下界更好的可行整数解;

(c) LPi 没有可行解.

分下面两种情况处理.

(a) 如果 LPi 已经探明并且找到了一个更好的解, 那么将这个值作为新的下界. 如果所有的子问题都已经探明, 停止; 整数线性规划的最优值就是当前这个有限的下界. 如果没有有限的下界, 那么这个问题就没有可行解. 否则, 令 $i=i+1$, 重复第 1 步.

(b) 如果 LPi 尚未探明, 那么转入第 2 步进行分支.

第 2 步 (分支) 在 LPi 的最优解中任意选择一个变量 x_j, 其最优值 x_j^* 不是整数. 删除区间

$$[x_j^*] < x_j < [x_j^*]+1$$

(其中 $[v]$ 表示不超过 v 的最大整数) 分别利用下面的两个约束生成两个线性规划子问题

$$x_j \leqslant [x_j^*] \text{ 和 } x_j \geqslant [x_j^*]+1$$

令 $i=i+1$, 转入第 1 步.

上面给出的步骤是求解最大化的问题. 对于最小化的情况, 只需要将下界用上界代替 (初始值令 $z=+\infty$).

B&B 算法可以很容易地扩展到求解混合整数规划问题 (这种问题中只有一部分变量是整数型的). 如果一个变量是连续型的, 在整个算法中将永远不会选择它作为分支变量. 在求解一个子问题之后, 如果在最优解中对应离散变量的值都是整数并且目标函数值改进了当前目标函数的界, 那么就用这个子问题最优值作为目标函数新的界.

习题 8.2A[1]

1. 应用 B&B 算法求解例 8.2-1, 以 x_2 作为分支变量. 从求解与 $x_2 \leqslant [x_2^*]$ 相关的子问题开始.

2. 写出下面每个问题的 B&B 树. 为简便起见, 开始总是选择 x_1 作为分支变量.

*(a)
$$\begin{aligned} \max\ & z = 3x_1 + 2x_2 \\ \text{s.t.}\ & 2x_1 + 5x_2 \leqslant 9 \\ & 4x_1 + 2x_2 \leqslant 9 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

(b)
$$\begin{aligned} \max\ & z = 2x_1 + 3x_2 \\ \text{s.t.}\ & 5x_1 + 7x_2 \leqslant 35 \\ & 4x_1 + 9x_2 \leqslant 36 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

(c)
$$\begin{aligned} \max\ & z = x_1 + x_2 \\ \text{s.t.}\ & 2x_1 + 5x_2 \leqslant 16 \\ & 6x_1 + 5x_2 \leqslant 27 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

*(d)
$$\begin{aligned} \min\ & z = 5x_1 + 4x_2 \\ \text{s.t.}\ & 3x_1 + 2x_2 \geqslant 5 \\ & 2x_1 + 3x_2 \geqslant 7 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

(e)
$$\begin{aligned} \max\ & z = 5x_1 + 7x_2 \\ \text{s.t.}\ & 2x_1 + x_2 \leqslant 13 \\ & 5x_1 + 9x_2 \leqslant 41 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

***3.** 假定 x_1 是连续的, 重新求解第 2 题.

4. 用图解法说明下面的整数线性规划没有可行解, 并利用 B&B 算法加以证明.

$$\begin{aligned} \max\ & z = 2x_1 + x_2 \\ \text{s.t.}\ & 10x_1 + 10x_2 \leqslant 9 \\ & 10x_1 + 5x_2 \geqslant 1 \\ & x_1, x_2 \geqslant 0\text{且均是整数} \end{aligned}$$

5. 利用 B&B 算法求解下面的问题.

$$\begin{aligned} \max\ & z = 18x_1 + 14x_2 + 8x_3 + 4x_4 \\ \text{s.t.}\ & 15x_1 + 12x_2 + 7x_3 + 4x_4 + x_5 \leqslant 37 \\ & x_1, x_2, x_3, x_4, x_5 = (0, 1) \end{aligned}$$

6. 将下面的问题转化成一个混合整数线性规划问题并求出最优解.

$$\begin{aligned} \max\ & z = x_1 + 2x_2 + 5x_3 \\ \text{s.t.}\ & |-x_1 + 10x_2 - 3x_3| \geqslant 15 \\ & 2x_1 + x_2 + x_3 \leqslant 10 \\ & x_1,\ x_2,\ x_3 \geqslant 0 \end{aligned}$$

7. TORA/Solver/AMPL 实验. 下面的问题用来论证即使问题的规模比较小, B&B 算法也可能需要计算较多的子问题. 特别要留意在找到最优解之前一共计算了多少个子问题, 并

[1] 这一节的习题都可以使用 AMPL、Excel 规划求解或者 TORA 的 MODIFY 选项, 通过交互式地求解子问题来得到上界或者下界.

且有多少个子问题被证明是最优的.

$$\begin{aligned} &\min \ y \\ &\text{s.t.} \ \ 2(x_1 + x_2 + \cdots + x_{15}) + y = 15 \\ &\qquad\qquad x_1, x_2, \cdots, x_{15}, y = (0, 1) \end{aligned}$$

(a) 用 TORA 的自动选项说明虽然最优解在第 9 个子问题时就找到了, 但是却检查了 25 000 多个子问题才最终确认.

(b) 应用 Excel 规划求解可以得到与 TORA 相同的结论. [注: 利用 Excel 规划求解, 可以从电子表格底部看到所产生分支 (子问题) 的数目的变化情况.]

(c) 应用 AMPL 求解这个问题, 可以立即得到最优解. 这么好的性能归功于 AMPL 执行了许多预备的步骤, 并且 CPLEX 具有求解这种问题的优势.

8. TORA 实验.考虑下面的整数线性规划:

$$\begin{aligned} &\max \ z = 18x_1 + 14x_2 + 8x_3 \\ &\text{s.t.} \ \ 15x_1 + 12x_2 + 7x_3 \leqslant 43 \\ &\qquad x_1,\ x_2,\ x_3 \text{均是非负整数} \end{aligned}$$

使用 TORA 的 B&B 用户向导选项, 在目标函数值的下界起作用和不起作用这两种情况下生成搜索树. 试分析使目标函数值的下界起作用对于产生的子问题数目有什么影响? 为了简便起见, 每次选择下标较小的变量作为分支变量, 并且首先按搜索树当前行从左到右的顺序选择子问题, 然后再考虑下一行.

***9.** TORA 实验.重新考虑上面的第 8 题. 将这个问题转化成等价的 $0-1$ 整数线性规划, 然后用 TORA 的自动选项求解. 试比较得到的两个搜索树的大小.

10. AMPL 实验.对于下面的 $0-1$ 整数线性规划, 用交互式的 AMPL 生成相应的搜索树. 在每一种情况下观察是如何利用 z 的界来探明子问题的.

$$\begin{aligned} &\max \ z = 3x_1 + 2x_2 - 5x_3 - 2x_4 + 3x_5 \\ &\text{s.t.} \quad x_1 + x_2 + x_3 + 2x_4 + x_5 \leqslant 4 \\ &\qquad 7x_1 + 3x_3 - 4x_4 + 3x_5 \leqslant 8 \\ &\qquad 11x_1 - 6x_2 + 3x_4 - 3x_5 \geqslant 3 \\ &\qquad x_1,\ x_2,\ x_3,\ x_4,\ x_5 = (0, 1) \end{aligned}$$

8.2.2 割平面算法

与 B&B 算法类似, 割平面算法也是从连续的线性规划最优解开始. 在连续模型的可行解空间内增加一些特殊的约束 (称为割), 并按照一定的方法来找到一个整数的最优极值点. 在例 8.2-2 中, 首先从图解的角度说明是如何利用割来得到整数解的, 然后给出代数学上的思想方法.

例 8.2-2

考虑下面的整数线性规划模型.

$$\begin{aligned} \max\ z &= 7x_1 + 10x_2 \\ \text{s.t.}\ -x_1 + 3x_2 &\leqslant 6 \\ 7x_1 + x_2 &\leqslant 35 \\ x_1, x_2 &\geqslant 0 \text{且均是整数} \end{aligned}$$

割平面算法就是通过增加产生了最优整数极值点的割, 来改变可行解空间. 图 8.10 给出了这样两个割的例子. 初始的时候, 得到了连续线性规划的最优解 $z = 66\frac{1}{2}, x_1 = 4\frac{1}{2}, x_2 = 3\frac{1}{2}$. 接下来, 增加割I, 可以得到连续线性规划的最优解 $z = 62, x_1 = 4\frac{4}{7}, x_2 = 3\frac{1}{2}$. 随后, 增加割II, 与割I一起增加到原始的约束中, 可以得到线性规划的最优解 $z = 58, x_1 = 4, x_2 = 3$, 这就是我们需要求的整数解.

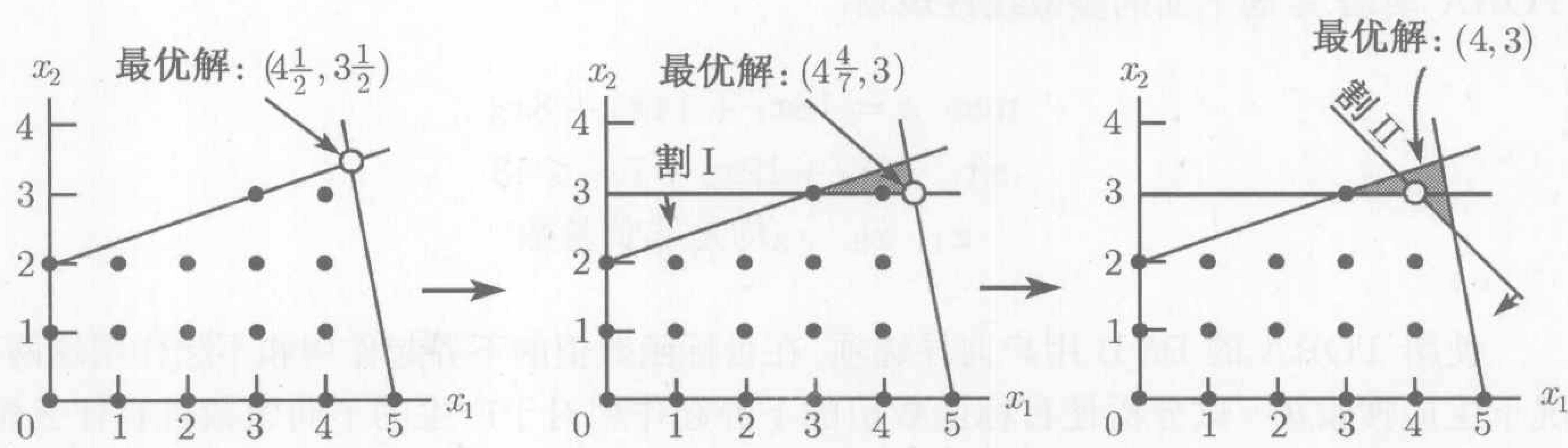


图 8.10　割在整数线性规划中的应用

增加的割保证不会删除任何一个原始问题的整数可行解, 同时割还必须穿过至少一个可行或者不可行的整数点. 这是割的最基本要求.

对于上面的例题, 两个变量的问题用了两个割就求得了最优解, 这也是一个巧合. 一般而言, 需要割的数目, 虽然是有限的, 但是它与问题的规模没有必然的联系. 也即是说, 一个变量个数和约束个数较少的问题需要割的数目, 可能会大于一个较大规模问题需要的割的数目.

接下来, 我们还是通过求解上面这个例题, 说明割是如何构造和使用的.

分别为约束 1 和 2 增加松弛变量 x_3 和 x_4. 最优的线性规划单纯形表是

基	x_1	x_2	x_3	x_4	解
z	0	0	$\frac{63}{22}$	$\frac{31}{22}$	$66\frac{1}{2}$
x_2	0	1	$\frac{7}{22}$	$\frac{1}{22}$	$3\frac{1}{2}$
x_1	1	0	$-\frac{1}{22}$	$\frac{3}{22}$	$4\frac{1}{2}$

最优解是 $z = 66\frac{1}{2}, x_1 = 4\frac{1}{2}, x_2 = 3\frac{1}{2}, x_3 = 0, x_4 = 0$. 割是建立在所有的变量 (包括松弛变量 x_3 和 x_4) 都是整数的假定之上的. 注意, 这个问题目标函数中所有的系数都是整数, 所以 z 的值也是整数.

最优单纯形表给出的信息可以写成下面的约束:

$$z+\tfrac{63}{22}x_3+\tfrac{31}{22}x_4=66\tfrac{1}{2}\quad(z\ \text{方程})$$
$$x_2+\tfrac{7}{22}x_3+\tfrac{1}{22}x_4=\ 3\tfrac{1}{2}\quad(x_2\ \text{方程})$$
$$x_1-\tfrac{1}{22}x_3+\tfrac{3}{22}x_4=\ 4\tfrac{1}{2}\quad(x_1\ \text{方程})$$

只要右端项是分数的约束方程都可以用作产生割的**原始行** (source row). 在这个例题中, 目标函数 z 恰好也是整数, 所以 z 方程也可以作为产生割的原始行. 下面我们将讨论从每一个原始行来产生割, 首先看 z 方程.

首先, 将方程中的每一个非整数系数分解成整数部分和分数部分, *保证分解出的分数部分是严格大于零的*. 例如,

$$\tfrac{5}{2}=\left(2+\tfrac{1}{2}\right)$$
$$-\tfrac{7}{3}=\left(-3+\tfrac{2}{3}\right)$$

按照这种分解方法, z 方程变为

$$z+\left(2+\tfrac{19}{22}\right)x_3+\left(1+\tfrac{9}{22}\right)x_4=\left(66+\tfrac{1}{2}\right)$$

将方程中所有整数部分移动到等式的左边, 所有的分数部分移动到等式的右边, 我们得到

$$z+2x_3+1x_4-66=-\tfrac{19}{22}x_3-\tfrac{9}{22}x_4+\tfrac{1}{2} \tag{1}$$

因为松弛变量 x_3 和 x_4 都是非负的, 并且所有原始的分数都是严格正的, 因此方程 (1) 的右端项一定满足下面的不等式:

$$-\tfrac{19}{22}x_3-\tfrac{9}{22}x_4+\tfrac{1}{2}\leqslant\tfrac{1}{2} \tag{2}$$

再者, 由于方程 (1) 左端项 $z+2x_3+x_4-66$ 一定是整数, 所以右端项 $-\frac{19}{22}x_3-\frac{9}{22}x_4+\frac{1}{2}$ 也一定是整数, 因此 (2) 式可以用下面的不等式代替:

$$-\tfrac{19}{22}x_3-\tfrac{9}{22}x_4+\tfrac{1}{2}\leqslant 0$$

满足 $\leqslant\frac{1}{2}$ 的整数值一定满足 $\leqslant 0$, 所以一定有上面的结果.

最后一个不等式就是所求的割, 它是得到整数解的一个*必要*(但不是充分的) 条件. 由于所有系数都是分数, 所以也称为**分数割** (fractional cut).

因为上面给出的连续线性规划问题的最优单纯形表中 $x_3=x_4=0$, 所以当前的连续解与割矛盾 (因为代入割得到 $\frac{1}{2}\leqslant 0$). 这样, 如果将割的约束增加到最后一张单纯形表中, 那么得到的最优解就会向满足整数要求的方向移动.

在说明如何在最优单纯形表中使用割之前, 我们先看看割是如何从其他约束方程构造的. 首先对于 x_1 行:

$$x_1 - \frac{1}{22}x_3 + \frac{3}{22}x_4 = 4\frac{1}{2}$$

按照上面的方法分解这个方程, 得到

$$x_1 + \left(-1 + \frac{21}{22}\right)x_3 + \left(0 + \frac{3}{22}\right)x_4 = \left(4 + \frac{1}{2}\right)$$

相应的割是

$$-\frac{21}{22}x_3 - \frac{3}{22}x_4 + \frac{1}{2} \leqslant 0$$

同理, 对于 x_2 方程:

$$x_2 + \frac{7}{22}x_3 + \frac{1}{22}x_4 = 3\frac{1}{2}$$

分解为

$$x_2 + \left(0 + \frac{7}{22}\right)x_3 + \left(0 + \frac{1}{22}\right)x_4 = 3 + \frac{1}{2}$$

因此, 相应的割可以写成

$$-\frac{7}{22}x_3 - \frac{1}{22}x_4 + \frac{1}{2} \leqslant 0$$

上面的 3 个割中的任何一个都可以在割平面算法的第一次迭代中使用. 因此一般情况下没有必要把 3 个割的表达式都写出来.

可以任意选择一个行来生成割, 比如 x_2 行, 把割写成等式的形式:

$$-\frac{7}{22}x_3 - \frac{1}{22}x_4 + s_1 = -\frac{1}{2}, \quad s_1 \geqslant 0 \quad (\text{割I})$$

将这个约束加到线性规划的最优单纯形表中, 如下:

基	x_1	x_2	x_3	x_4	s_1	解
z	0	0	$\frac{63}{22}$	$\frac{31}{22}$	0	$66\frac{1}{2}$
x_2	0	1	$\frac{7}{22}$	$\frac{1}{22}$	0	$3\frac{1}{2}$
x_1	1	0	$-\frac{1}{22}$	$\frac{3}{22}$	0	$4\frac{1}{2}$
s_1	0	0	$-\frac{7}{22}$	$-\frac{1}{22}$	1	$-\frac{1}{2}$

这个单纯形表是最优不可行的, 所以使用对偶单纯形法 (4.4.1 节) 来恢复可行性, 得到

基	x_1	x_2	x_3	x_4	s_1	解
z	0	0	0	1	9	62
x_2	0	1	0	0	1	3
x_1	1	0	0	$\frac{1}{7}$	$-\frac{1}{7}$	$4\frac{4}{7}$
x_3	0	0	1	$\frac{1}{7}$	$-\frac{22}{7}$	$1\frac{4}{7}$

最优解中 x_1 和 x_3 仍然不是整数. 任意选择 x_1 作为原始行, 那么有

$$x_1 + \left(0 + \tfrac{1}{7}\right) x_4 + \left(-1 + \tfrac{6}{7}\right) s_1 = 4 + \tfrac{4}{7}$$

相应的割为

$$-\tfrac{1}{7}x_4 - \tfrac{6}{7}s_1 + s_2 = -\tfrac{4}{7}, \quad s_2 \geqslant 0 \quad (\text{割II})$$

基	x_1	x_2	x_3	x_4	s_1	s_2	解
z	0	0	0	1	9	0	62
x_2	0	1	0	0	1	0	3
x_1	1	0	0	$\frac{1}{7}$	$-\frac{1}{7}$	0	$4\frac{4}{7}$
x_3	0	0	1	$\frac{1}{7}$	$-\frac{22}{7}$	0	$1\frac{4}{7}$
s_2	0	0	0	$-\frac{1}{7}$	$-\frac{6}{7}$	1	$-\frac{4}{7}$

应用对偶单纯形法得到下表:

基	x_1	x_2	x_3	x_4	s_1	s_2	解
z	0	0	0	0	3	7	58
x_2	0	1	0	0	1	0	3
x_1	1	0	0	0	−1	1	4
x_3	0	0	1	0	−4	1	1
x_4	0	0	0	1	6	−7	4

这样就得到了所求的整数解 ($x_1 = 4, x_2 = 3, z = 58$). 注意, 最后一张单纯形表中的所有系数都是整数, 这并不是巧合, 而是执行分数割过程中的一个固有性质决定的.

评注 在使用分数割时一条非常重要的假设就是, 所有的变量 (包括松弛变量和剩余变量) 都是整数型的. 这就意味着割只能处理纯整数问题. 可以通过下面一个例子来说明这条假设的重要性.

考虑约束

$$x_1 + \tfrac{1}{3}x_2 \leqslant \tfrac{13}{2}$$

$$x_1, x_2 \geqslant 0 \text{且均是整数}$$

从求解相应整数线性规划的观点出发, 通过增加非负松弛变量 s_1 可以将这个约束化为下面等式:

$$x_1+\frac{1}{3}x_2+s_1=\frac{13}{2}$$

应用分数割时假定这个约束有一个满足所有变量 x_1, x_2, s_1 都是整数的可行解. 然而从上面的等式可以看出, 只有当 s_1 是非整数的时候, 才可能有满足 x_1 和 x_2 都为整数的可行解. 这也就说明, 如果应用割平面算法将得不到可行的整数解, 但事实上在原始问题上 x_1 和 x_2 是有可行整数解的.

对于这种情况, 有下面两种弥补方法.

(1) 在原始的约束两边乘以一个适当的常数, 使得所有的相关系数都变为整数. 例如, 在上面的例子中两边同时乘以 6, 可以得到

$$6x_1+2x_2\leqslant 39$$

这样任何的整数解 x_1 和 x_2 都会自动地有一个整数的松弛. 然而, 这种转化方法只适用于简单的约束, 对于一些复杂的情形, 为了整数化系数可能会乘以一个非常大的常数.

(2) 应用一种特殊的割, 称作**混合割** (mixed cut), 它允许只有一部分变量取整数值. 其他的变量 (包括松弛变量和剩余变量) 可以是连续的变量. 本章将不再详细介绍这种割 (见 Taha, 1975, pp.198-202).

习题 8.2B

1. 在例 8.2-2 中, 用图解法说明下面的约束是否可以作为可行的割:

***(a)** $x_1+2x_2\leqslant 10$　　**(b)** $2x_1+x_2\leqslant 10$

(c) $3x_2\leqslant 10$　　**(d)** $3x_1+x_2\leqslant 15$

2. 在例 8.2-2 中, 用图解法说明应用下面的两个 (可行的) 割是如何得到最优整数解的:

$$x_1+2x_2\leqslant 10\quad(\text{割I})$$
$$3x_1+x_2\leqslant 15\quad(\text{割II})$$

3. 在例 8.2-2 中, 分别按照 x_1 和 x_2 表示出割 I 和 II, 并从图 8.10 中说明它们是相同的.

4. 在例 8.2-2 中, 按照 x_3 行产生割 II, 试继续求解.

5. 试说明对于下面的问题, 虽然它存在可行的整数解 x_1 和 x_2, 但是使用分数割来处理就求不出可行解, 除非将约束中的分数系数转化为整数系数.

$$\max\ z=x_1+2x_2$$
$$\text{s.t.}\ \ x_1+\frac{1}{2}x_2\leqslant\frac{13}{4}$$
$$x_1, x_2\geqslant 0\text{且均是整数}$$

6. 应用分数割求解下面的问题, 并且对得出的最优解与将连续最优解舍入得到的整数解进行比较.

*(a) $\max\ z = 4x_1 + 6x_2 + 2x_3$

$$\begin{aligned} \text{s.t.}\quad & 4x_1 - 4x_2 \leqslant 5 \\ & -x_1 + 6x_2 \leqslant 5 \\ & -x_1 + x_2 + x_3 \leqslant 5 \\ & x_1, x_2, x_3 \geqslant 0 \text{且均是整数} \end{aligned}$$

(b) $\max\ z = 3x_1 + x_2 + 3x_3$

$$\begin{aligned} \text{s.t.}\quad & -x_1 + 2x_2 + x_3 \leqslant 4 \\ & 4x_2 - 3x_3 \leqslant 2 \\ & x_1 - 3x_2 + 2x_3 \leqslant 3 \\ & x_1, x_2, x_3 \geqslant 0 \text{且均是整数} \end{aligned}$$

8.2.3 整数线性规划的计算性分析

尽管已经研究了 40 余年, 但到目前为止仍然没有一个统一的计算机软件可以求解所有的整数线性规划问题. 不过, 在本章所讲述的两种算法中, B&B 算法更可靠一些, 实际上, 所有的求解整数线性规划的商用软件都是基于 B&B 算法的. 割平面算法一般比较复杂而且不可靠, 其中的舍入误差也会带来很大的麻烦. 因为计算机中存储分数的精确性决定了割的精确性. 例如, 在例 8.2-2 中, 计算机不可能精确地用浮点型的小数来表示分数 $\frac{1}{7}$, 不管所用的精度有多高. 虽然已经做了许多工作来改进割平面算法的计算效率, 但是目前仍然没有太大的进展. 在更多的情况下, 割平面算法是作为 B&B 算法的补充, 对 B&B 算法中的每个子问题使用割平面算法来减少一部分解空间.

影响整数规划计算性的一个重要因素, 就是整数变量的个数以及它们相应取值的可行区间. 由于一般可用的算法所能求解整数线性规划的数目有限, 所以尽量地减少整数线性规划中的整数变量的个数, 这对我们求解一个整数线性规划问题会有很大帮助. 下面的 3 条技巧会经常用到:

(1) 如果可能的话, 尽量用一个连续的变量来近似代替一个整数变量;

(2) 对于一个整数变量, 尽量缩小它的可行区间;

(3) 在模型中尽量避免使用非线性的约束和目标函数.

整数规划在现实中有十分重要的应用, 然而求解整数规划的可靠算法却很难满足我们的需要. 由于离散数学问题的性质以及整数可行解区域不是凸集, 这使得研究人员不大可能在整数规划的理论上有新的突破. 相反, 计算机 (软件和硬件) 新技术的发展也带来了改进整数线性规划代码的希望.

8.3 旅行商问题 (TSP)

经典的 TSP 问题是在 n 个城市的情形下寻找一条经过每个城市恰好一次的最短的 (闭) 游程. 实质上, 这个问题是一个避免子游程的指派模型. 在 n 个城市的特定情形下, 定义

$$x_{ij}=\begin{cases}1, & \text{如果从城市 } i \text{ 到城市 } j\\ 0, & \text{如果从城市 } i \text{ 不能到达城市 } j\end{cases}$$

已知从城市 i 到城市 j 的距离是 d_{ij}, 那么 TSP 模型可以写成

$$\min \quad z=\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij}, \quad \text{如果 } i=j, d_{ij}=\infty$$

$$\text{s.t.} \quad \sum_{j=1}^{n} x_{ij}=1, \quad i=1,2,\cdots,n \tag{1}$$

$$\sum_{i=1}^{n} x_{ij}=1, \quad j=1,2,\cdots,n \tag{2}$$

$$x_{ij} \quad =(0,1) \tag{3}$$

$$\text{解构成了一个 } n \text{ 个城市的游程} \tag{4}$$

约束 (1), (2), (3) 定义了一个标准的指派问题 (5.4 节). 图 8.11 给出了一个 5 个城市的例子, 每一条边都是双向连通的. 图中还给出了可以相应指派问题的游程解和一个子游程解. 不考虑约束 (4), 如果指派问题的最优解刚好对应一个游程, 那么这个游程就是 TSP 的最优解. 约束 (4) 就是为了保证得到的解是一个游程.

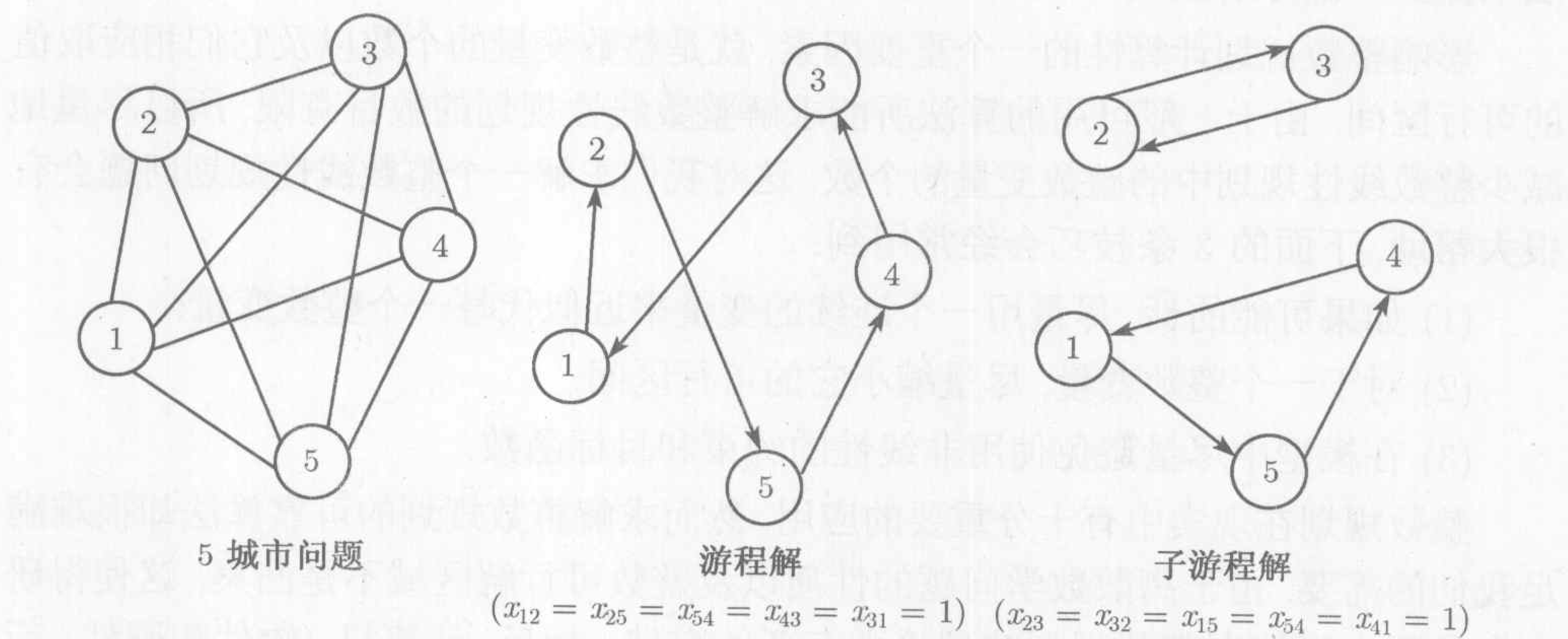


图 8.11　5 个城市的 TSP 实例, 相应的指派问题模型可以给出游程解和子游程解

求解 TSP 问题的精确算法有分支限界算法和割平面算法. 这两种算法都是基于 8.2 节给出的一般 B&B 算法和割平面算法的思想. 然而这个问题是一个典型的难解的问题, 也即是求解这个问题所需要的存储空间和计算时间都非常大. 正是由于这个原因, 在某些时候一些启发式的算法也可以给出问题比较“好”的解.

在给出启发式算法和精确算法之前, 我们先来通过一个例题说明许多实际问题都可以转化为 TSP 模型. (见习题 8.3A)

例 8.3-1

彩虹公司每天都需要生产一批颜料, 包含白色 (W)、黄色 (Y)、红色 (R) 和黑色 (B), 下面需要安排一种生产加工 4 种颜料的顺序. 由于彩虹公司使用同一套设备生产这 4 种颜色的颜料, 所以在生产另外一种颜色的颜料之前必须对设备进行清洗. 下表给出了生产下一种颜色之前需要清洗设备的时间. 由于一种颜色的颜料是按批量生产的, 所以在表格中对角线上的时间定义为无穷. 目标就是要寻找一个加工 4 种颜色颜料的顺序, 使得总的清洗时间最少.

	生产下一种颜料之前的清洗时间 (分钟)			
当前颜料	白色	黄色	黑色	红色
白色	∞	10	17	15
黄色	20	∞	19	18
黑色	50	44	∞	25
红色	45	40	20	∞

每一种颜料看作一个"城市", 从加工一种颜料到另一种颜料需要的清洗时间表示两个"城市"之间的"距离", 那么这个问题就可以转化成求一条**最短的回路**, 使得从一种颜料出发经过其他 3 种颜料恰好一次, 然后回到初始的颜料.

对于这个问题, 我们使用枚举的方法求解, 这个网络一共有 6 种 ($(4-1)! = 3! = 6$) 可能的回路. 从下表中可以看出最优的回路是 $W \to Y \to R \to B \to W$.

生产的游程	总的清洗时间
$W \to Y \to B \to R \to W$	$10+19+25+45=99$
$W \to Y \to R \to B \to W$	$10+18+20+50=98$
$W \to B \to Y \to R \to W$	$17+44+18+45=124$
$W \to B \to R \to Y \to W$	$17+25+40+20=102$
$W \to R \to B \to Y \to W$	$15+20+44+20=99$
$W \to R \to Y \to B \to W$	$15+40+19+50=124$

但是一般情况下对游程进行枚举的方法是不切实际的. 例如对于一个 11 个城市的问题, 如果用枚举的方法, 那么会有 $10! = 3\ 628\ 800$ 种游程, 检查每一个游程将是一个非常庞大的计算量. 正是由于这个原因, 需要改进问题的模型, 并引进新的方法求解, 这将在下面介绍.

为该颜料问题建立基于指派的模型, 我们定义

$$x_{ij} = \begin{cases} 1, & \text{颜料 } i \text{ 在颜料 } j \text{ 之前生产} \\ 0, & \text{颜料 } i \text{ 在颜料 } j \text{ 之后生产} \end{cases}$$

取 M 是一个足够大的正数, 那么彩虹公司的问题可以建立以下模型:

$$
\begin{aligned}
\min\ z = & Mx_{WW} + 10x_{WY} + 17x_{WB} + 15x_{WR} + 20x_{YW} + Mx_{YY} \\
& + 19x_{YB} + 18x_{YR} + 50x_{BW} + 44x_{BY} + Mx_{BB} + 25x_{BR} \\
& + 45x_{RW} + 40x_{RY} + 20x_{RB} + Mx_{RR}
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t. } & x_{WW} + x_{WY} + x_{WB} + x_{WR} = 1 \\
& x_{YW} + x_{YY} + x_{YB} + x_{YR} = 1 \\
& x_{BW} + x_{BY} + x_{BB} + x_{BR} = 1 \\
& x_{RW} + x_{RY} + x_{RB} + x_{RR} = 1 \\
& x_{WW} + x_{YW} + x_{BW} + x_{RW} = 1 \\
& x_{WY} + x_{YY} + x_{BY} + x_{RY} = 1 \\
& x_{WB} + x_{YB} + x_{BB} + x_{RB} = 1 \\
& x_{WR} + x_{YR} + x_{BR} + x_{RR} = 1 \\
& x_{ij} = (0, 1) \text{ 对所有的 } i \text{ 和 } j \\
& \text{解是一个游程 (回路)}
\end{aligned}
$$

在目标函数中使用 M 是为了避免一种颜料生产之后经过清洗设备再次生产它, 同样也可以将整个模型中的变量 $x_{WW}, x_{YY}, x_{BB}, x_{RR}$ 删除.

习题 8.3A

*1. 一位经理雇用了 10 个雇员来做 6 个项目. 下表给出了雇员所做的项目情况:

雇员 \ 项目	1	2	3	4	5	6
1		x		x	x	
2	x		x		x	
3		x	x	x		x
4			x	x	x	
5	x	x	x			
6	x	x	x	x		x
7	x	x			x	x
8	x		x	x		
9					x	x
10	x	x		x	x	x

经理每周都来会见每一个雇员听取进度报告. 会见 1 个雇员需要 20 分钟, 10 个雇员总共需要 3 小时 20 分钟. 为了节约时间, 经理决定按项目来召开会议, 目标是将需要召开的会议排一个顺序, 使得总的出入会议室的人数尽量少. 根据这个问题建立数学模型.

2. 居住在 Basin 的一位售书商每月需要到 4 个分别位于 Wald、Bon、Mena 和 Kiln 的客户一次. 下表给出了几个城市之间的相互距离, 单位是英里.

	城市之间的距离 (英里)				
	Basin	Wald	Bon	Mena	Kiln
Basin	0	120	220	150	210
Wald	120	0	80	110	130
Bon	220	80	0	160	185
Mena	150	110	160	0	190
Kiln	210	130	185	190	0

目标就是使得售书商需要走的总路程最小. 将这个问题建立成基于指派的整数线性规划模型.

3. 电路板 (例如个人电脑上的电路板) 上需要打许多的孔用来安装各种电子元件. 这些孔是用一个可移动的钻孔机来钻孔的. 下表给出了一个特殊的电路板上 6 个需要钻的孔之间的距离.

$$\|d_{ij}\| = \begin{pmatrix} - & 1.2 & 0.5 & 2.6 & 4.1 & 3.2 \\ 1.2 & - & 3.4 & 4.6 & 2.9 & 5.2 \\ 0.5 & 3.4 & - & 3.5 & 4.6 & 6.2 \\ 2.6 & 4.6 & 3.5 & - & 3.8 & 0.9 \\ 4.1 & 2.9 & 4.6 & 3.8 & - & 1.9 \\ 3.2 & 5.2 & 6.2 & 0.9 & 1.9 & - \end{pmatrix}$$

用一个整数线性规划模型中的指派部分来表示这个问题.

8.3.1 启发式算法

本节介绍两种启发式算法: *最近邻域和颠倒子游程算法*. 第 1 种算法执行起来很容易, 第 2 种算法相对复杂一些, 但是第 2 种算法得到的解比第 1 种算法要好. 最后, 我们将这两个算法合并得到一个启发式算法, 也就是, 最近邻域算法的输出用作颠倒子游程算法的输入.

最近邻域启发式算法 从这个算法的名字就可以看出算法的原理是什么, 为了得到 TSP 问题的一个较"好"的解, 我们从任意的一个城市出发, 选择最近的城市连接, 然后新加进来的城市在没有到达过的城市中选择一个最近的连接起来, 这样一直连接下去直到形成一个游程.

例 8.3-2

下面的表格给出了一个 5 个城市 TSP 问题的城市之间的距离 (英里).

$$\|d_{ij}\| = \begin{pmatrix} \infty & 120 & 220 & 150 & 210 \\ 120 & \infty & 100 & 110 & 130 \\ 220 & 80 & \infty & 160 & 185 \\ 150 & \infty & 160 & \infty & 190 \\ 210 & 130 & 185 & \infty & \infty \end{pmatrix}$$

启发式算法可以从任何一个城市出发, 但是从不同的城市出发得到的游程可能也是不同的. 下面的表格给出了从城市 3 出发的启发式算法的运行过程.

步	执行的动作	(部分的) 游程
1	从城市 3 开始	3
2	连接城市 2, 因为它是离城市 3 最近的 ($d_{32} = \min\{220, \mathbf{80}, \infty, 160, 185\}$)	$3-2$
3	连接城市 4, 因为它是离城市 2 最近的 ($d_{24} = \min\{120, \infty, —, \mathbf{110}, 130\}$)	$3-2-4$
4	连接城市 1, 因为它是离城市 4 最近的 ($d_{41} = \min\{\mathbf{150}, \infty, —, —, 190\}$)	$3-2-4-1$
5	自动连接城市 5, 再连接回到城市 3	$3-2-4-1-5-3$

注意到, 在算法执行的每一步中, 比较距离的时候我们不必要考虑已经连接到游程上的城市. 在表格的中间列中, 我们使用 (—) 表示不需要考虑的距离.

最终得到的游程 $3-2-4-1-5-3$ 的长度是 $80+110+150+210+185=735$ 英里. 应该注意的是, 如果从不同的城市出发, 则可能得到不同的游程, 所以这种启发式算法依赖于所选取的出发城市. 例如, 在这个例题中如果选择城市 1 开始, 那么最终的游程是 $1-2-3-4-5-1$, 长度是 780 英里.

颠倒子游程启发式算法　对于 n 个城市的情形, 颠倒子游程算法从一个已经得到的可行的游程开始, 首先通过颠倒其中的任意 2 个城市的子游程来改进初始的游程, 然后再考虑 3 个城市的子游程, 直到检查所有的长度为 $(n-1)$ 的子游程.

例 8.3-3

我们仍然考虑例 8.3-2 给出的问题, 从长度为 745 英里的可行游程 $1-4-3-5-2-1$ 开始, 下表给出了执行的颠倒步骤:

类型	颠倒	游程	长度
开始	—	$(1-4-3-5-2-1)$	**745**
每次颠倒两个城市	$4-3$	$1-\underline{3-4}-5-2-1$	820
	$3-5$	$(1-4-\underline{5-3}-2-1)$	**725**
	$5-2$	$1-4-3-\underline{2-5}-1$	730
每次颠倒 3 个城市	$4-5-3$	$1-\underline{3-5-4}-2-1$	∞
	$5-3-2$	$1-4-\underline{2-3-5}-1$	∞
每次颠倒 4 个城市	$4-5-3-2$	$1-\underline{2-3-5-4}-1$	∞

以 $1-4-3-5-2-1$ 为初始游程, 每次颠倒两个城市得到 $4-3$, $3-5$ 和 $5-2$, 对应得到的游程的长度分别为 820, 725 和 730, 因为游程 $1-4-5-3-2-1$ 的长度较小 (= 725), 所以取这个游程作为颠倒 3 个城市的初始游程. 正像表格中给出的, 每

次颠倒 3 个城市没有能够产生较好的游程, 所以仍然使用它作为每次颠倒 4 个城市的初始游程, 依此类推, 最终按照这种启发式算法得到的游程是 $1-4-5-3-2-1$, 长度为 725 英里.

注意到, 在每次颠倒 3 个城市的步骤中没有能够得到更好的游程, 因此在每次颠倒 4 个城市的步骤中仍然使用上一次的游程. 又注意到我们在进行两个城市颠倒的时候没有考虑颠倒初始城市 (在这个例题中就是城市 1), 这是因为如果颠倒的话就不是一个游程了. 例如, 颠倒 $1-4$ 得到的 $4-1-3-5-3-2-1$ 就不是一个游程.

根据颠倒子游程启发式算法得到的解与初始游程有很大关系. 例如, 如果从长度为 750 英里的游程 2–3–4–1–5–2 开始, 那么最终得到的游程是 2–1–4–3–5–2, 长度是 745 英里, 这个解比上表中得到的解要差. 正是由于这个原因, 我们可以首先使用最近邻域算法求出从每个城市出发的所有游程, 然后取一个最好的游程作为颠倒子游程启发式算法的初始游程. 显然如果把这两种启发式算法结合起来就可以得到比分别使用时更好的解. 下表中给出了结合这两种启发式算法求解当前例子的过程.

启发式算法	初始城市	游程	长度
	1	$1-2-3-4-5-1$	780
	2	$2-3-4-1-5-2$	750
最近邻域	3	**$(3-2-4-1-5-3)$**	**735**
	4	$4-1-2-3-5-4$	∞
	5	$5-2-3-4-1-5$	750
	$2-4$	$3-\underline{4-2}-1-5-3$	∞
	$4-1$	$(3-2-\underline{1-4}-5-3)$	**725**
	$1-5$	$3-2-4-\underline{5-1}-3$	810
颠倒子游程	$2-1-4$	$3-\underline{4-1-2}-5-3$	745
	$1-4-5$	$3-2-\underline{5-4-1}-3$	∞
	$2-1-4-5$	$3-\underline{5-4-1-2}-3$	∞

Excel 程序

图 8.12 给出了启发式算法的一般 Excel 模板 (文件 excelTSP.xls). 在 H3 中有 3 个命令选项:

(1) 如果在其中输入一个指定城市, 那么就会运行最近邻域算法来从指定的城市开始求解.

(2) 如果输入"tour"(不含引号), 同时在指定空间里输入一个初始可行游程, 那么就会以这个游程开始应用颠倒启发式算法求解.

(3) 如果输入"all", 那么执行过程先应用最近邻域算法, 然后根据得到的最好

的游程来执行颠倒启发式算法.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1						Traveling Salesperson Nearest-Neighbor + Subtour Reversal Heuristics									
2	*Input steps:* (See commnet cell A4)				*Output steps:* (See comment cell H3)				*Distance matrix:* (infinity = blank cell)						
3	*Step 1:*	Number of cities =		5	*Step 3:*	Start heuristic with:		all		1	2	3	4	5	
4	*Step 2:*	Click to enter input data			*Step 4:*	Click to execute heuristic			1		120	220	150	210	
5									2	120		100	110	130	
6	*Solution sumamry:*								3	220	80		160	185	
7	Start city	Tour						Length	4	150		160		190	
8	1	1-2-3-4-5-1						780	5	210	130	185			
9	2	2-3-4-1-5-2						750		Initial tour:					
10	3	3-2-4-1-5-3						735		1	2	3	4	5	1
11	4	4-1-2-3-5-4						infinity							
12	5	5-2-3-4-1-5						750							
13	Reversals														
14	2-4	3-4-2-1-5-3						infinity							
15	4-1	3-2-1-4-5-3						725							
16	1-5	3-2-4-5-1-3						810							
17															
18	2-1-4	3-4-1-2-5-3						745							
19	1-4-5	3-2-5-4-1-3						infinity							
20															
21	2-1-4-5	3-5-4-1-2-3						infinity							

图 8.12　利用 Excel 电子表格执行 TSP 启发式算法 (文件 excelTSP.xls)

文件 excelTSP.xls 自动按第 3 种操作选项执行.

习题 8.3B

1. 应用启发式算法求解下面的问题:

   (a) 例 8.3-1 中的颜料生产顺序问题.　(b) 习题 8.3A 中的第 1 题.

   (c) 习题 8.3A 中的第 2 题.　(d) 习题 8.3A 中的第 3 题.

8.3.2　B&B 算法

B&B 算法的思想就是首先从相应的指派问题的最优解出发, 如果得到的解是一个游程, 那么求解过程结束. 否则, 通过增加限制约束来处理子游程, 可以通过创建与子游程中变量 x_{ij} 数目相同的分支来实现. 对应每一个分支的约束就是将这个子游程中的这一个变量设为 0(因为子游程中相应的变量值都是 1). TSP 问题相应的指派问题有可能产生一个游程, 也可能产生不了游程. 如果产生了游程, 我们用这个游程的目标值作为最小游程的一个上界. 如果产生不了游程, 则需要继续进行分支, 对于一个子游程需要产生的分支数目是子游程中变量的数目. 依次进行, 直到所有的分支都已经探明, 即要么这个分支得到了一个更好的 (更小的) 上界, 要么很明显这个分支不能够得到一个更好的游程. 最优的游程就是对应于最好的上界的游程.

下面的例题仔细讲解了关于 TSP 的 B&B 算法.

例 8.3-4

考虑下面的 5 个城市的 TSP 问题:

$$\|d_{ij}\| = \begin{pmatrix} \infty & 10 & 3 & 6 & 9 \\ 5 & \infty & 5 & 4 & 2 \\ 4 & 9 & \infty & 7 & 8 \\ 7 & 1 & 3 & \infty & 4 \\ 3 & 2 & 6 & 5 & \infty \end{pmatrix}$$

从求解这个问题相应的指派问题出发, 可以得到下面的解:

$$z = 15, x_{13} = x_{31} = 1, x_{25} = x_{54} = x_{42} = 1, \quad \text{其他的变量} = 0$$

这个解给出了两个子游程: $(1-3-1)$ 和 $(2-5-4-2)$. 图 8.13 的节点 1 处表示出了这两个子游程. 相应的总长度为 $z = 15$, 这个值是上述 TSP 问题的一个下界.

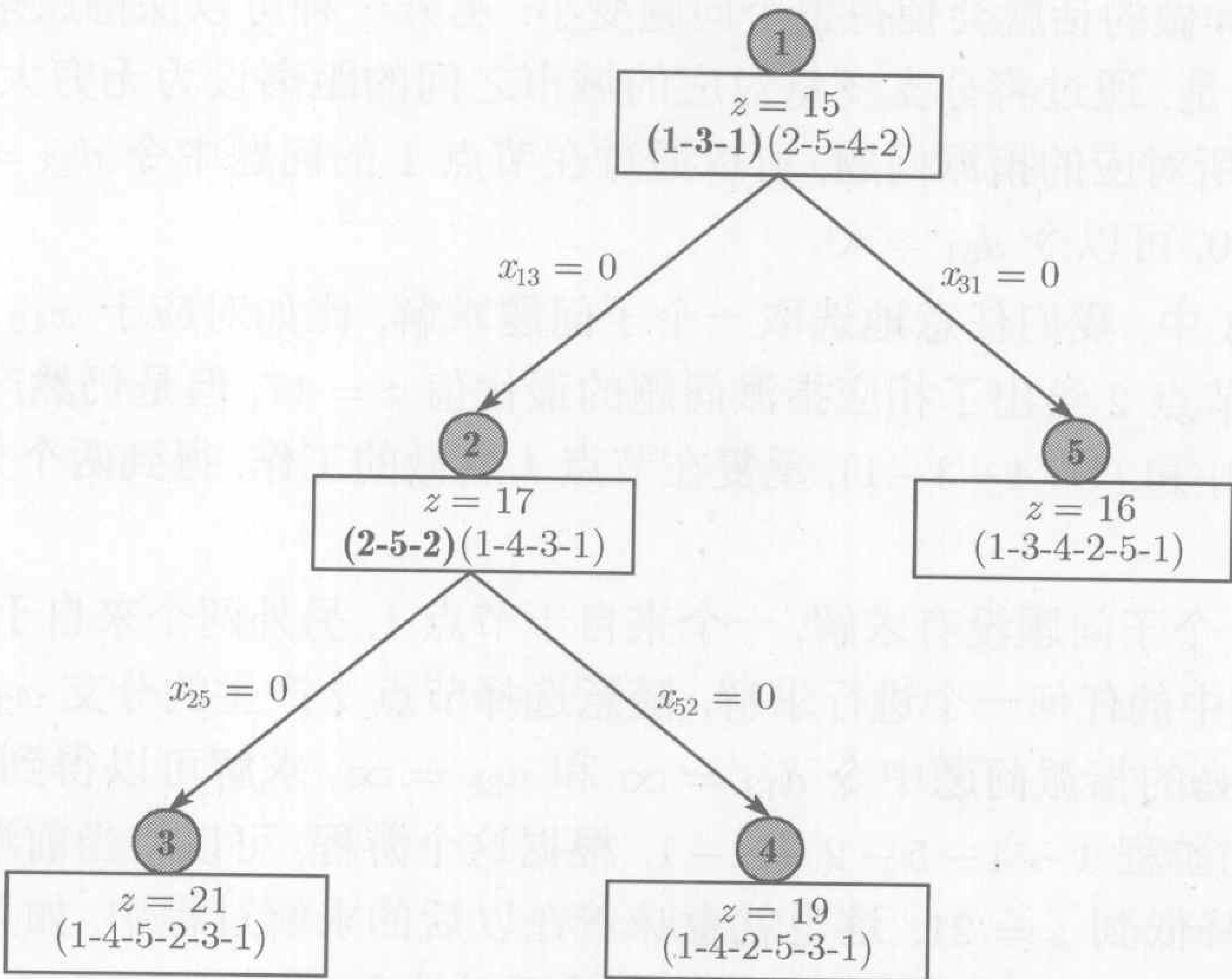


图 8.13　例 8.3-4 中利用 B&B 算法求解 TSP 问题的过程

当然我们很容易确定最优解的一个上界, 比如任意选取一个游程的长度作为最优游程的一个上界估计, 例如任意选取游程 $1-2-3-4-5-1$, 总长度为 $10+5+7+4+3 = 29$. 当然也可以使用 8.3.1 节中给出的启发式算法得到比较好的一个上界. 在这里我们就选取上面这个长度为 29 的游程作为使用 B&B 算法的一个上界. 然后, 我们使用启发式算法给出的较好的上界, 来看看它对搜索树有什么影响.

根据上面提到的上界和下界, 可以知道最优的游程的长度一定介于 (15,29). 如果求解过程中得到一个游程的长度大于或者等于 29, 那么这个解就可以舍弃掉.

为了处理节点 1 上的子游程, 我们需要将相应的回路破坏掉, 方法就是令组成这个子游程的变量 x_{ij} 等于 0. 例如, 对于子游程 $1-3-1$, 可以通过在节点 1 的指派问题中增加约束 $x_{13}=0$ 或者 $x_{31}=0$(一个分支只需要选择其中之一) 来破坏这个子游程, 同样地对于子游程 $2-5-4-2$, 也可以增加相应的约束 $x_{25}=0$、$x_{54}=0$ 或者 $x_{42}=0$ 来破坏这个子游程. 对于 B&B 树而言, 每增加一个这样的约束就相应得到一个新的子问题. 值得注意的一点就是在节点 1 上没有必要对两个子游程都进行破坏, 而是在每一个节点只需要对其中的一个子游程产生分支. 这是因为, 对其中之一进行破坏就会自动地调整其他子游程中的变量, 也就可以提供产生游程的条件了. 在这种假设条件下, 选择城市个数最少的子游程会对求解更加高效, 因为这样的话得到的分支数就会最少.

对于这个问题, 选取子游程 $(1-3-1)$, 那么在节点 1 就产生了两个分支 $x_{13}=0$ 和 $x_{31}=0$. 对于每个分支的指派问题, 只需要将对应取 0 的变量所在的行和列删除掉, 显然这样做的话就会使得整个问题变小. 另外一种可以保持原来指派问题规模不变的方法是, 通过将分支变量对应的城市之间的距离设为无穷大. 例如, 增加约束 $x_{13}=0$ 所对应的指派问题, 可以通过在节点 1 的问题中令 $d_{13}=\infty$ 得到; 同样对于 $x_{31}=0$, 可以令 $d_{31}=\infty$.

在图 8.13 中, 我们任意地选取一个子问题求解, 比如对应于 $x_{13}=0$ 的分支, 令 $d_{13}=\infty$. 节点 2 给出了相应指派问题的最优值 $z=17$, 但是仍然产生了两个子游程 $(2-5-2)$ 和 $(1-4-3-1)$, 重复在节点 1 所做的工作, 得到两个分支: $x_{25}=0$ 和 $x_{52}=0$.

现在有 3 个子问题没有求解, 一个来自于节点 1, 另外两个来自于节点 2, 当然可以选择它们中的任何一个进行求解. 随意选择节点 2 产生的分支 $x_{25}=0$ 对应的子问题, 在*原始*的指派问题中令 $d_{13}=\infty$ 和 $d_{25}=\infty$, 求解可以得到 $z=21$ 以及节点 3 相应的游程 $1-4-5-2-3-1$. 根据这个游程, 可以将当前游程长度的上界从 $z=29$ 降低到 $z=21$. 这也就意味着在以后的求解过程中, 如果得到了一个游程对应的长度大于 21, 那么相应的分支就可以被舍弃.

现在还有两个问题没有考虑. 随意选择子问题 4, 在*原始*问题中令 $d_{13}=\infty$ 和 $d_{52}=\infty$, 然后求解得到游程 $1-4-2-5-3-1$, 相应的长度是 $z=19$, 这个解提供了一个新的游程, 比当前游程长度的上界 21 更小, 因此可以将上界改进到 $z=19$, 对应的游程是 $1-4-2-5-3-1$, 这也是目前为止最好的解.

下面只剩下子问题 5 没有考虑了. 在*原始*问题中令 $d_{31}=\infty$, 求解可以得到游程 $1-3-4-2-5-1$, 长度 $z=16$. 同样这个解又比节点 4 给出的游程好, 因此最优游程长度的上界更改为 $z=16$.

现在所有的子问题都已经考虑过了, 也就完成了搜索树的构造, 最终得到的最优游程就是当前的上界所对应的游程 $1-3-4-2-5-1$, 长度是 16 英里.

评注 这个例题的解可以揭示出下面的两点:

(1) 虽然在搜索树中选择的节点序列 $1 \to 2 \to 3 \to 4 \to 5$ 是为了故意来演示 B&B 算法的求解过程以及上界的变更情况, 但是一般情况下, 我们没有办法预言到底在搜索过程中哪一个序列更加高效, 不过我们仍然有些选取的规则对求解有一定的帮助. 例如, 在一个给定的节点处, 一般选择分支中对应d_{ij} 最大的那一支, 这样做就是通过将形成游程的最长的边去掉, 从而有望得到一个"好"的游程. 在上面的例题中, 按照这个规则, 我们在节点 1 处会首先选择节点 5(而不是节点 2) 求解, 因为 $(d_{31} = 4) > (d_{13} = 3)$, 所以选取 $x_{31} = 0$ 对应的一支, 求解这一支就可以得到新的上界 $z = 16$, 那么节点 2 就不需要再进行分支了, 因而也就不需要求解节点 3 和节点 4. 另外一条规则就是, 在考察节点序列的时候, 按照首先考虑水平方向 (而不是竖直方向), 也就是优先考虑距离原始节点近的节点, 因为在这样的节点上增加的约束数目少, 更容易产生好的上界. 按照这个规则, 节点 5 也会比节点 3 和节点 4 被更快地考虑到.

(2) 一般 B&B 算法应该与 8.3.1 节中的启发式算法结合起来使用. 启发式算法可以给出一个"好"的上界, 从而在搜索树中能够舍弃掉更多的分支. 在上面的例题中, 按照启发式算法就可以得到游程 $1-3-4-2-5-1$, 长度为 16.

AMPL 程序

交互式的 AMPL 命令非常适合于应用了一般指派问题模型来实现 TSP 问题的 B&B 算法 (文件 amplAssignment.txt). 下表总结了在图 8.13 中创建 B&B 树所需要的 AMPL 命令 (例 8.3-4):

AMPL 命令	结　果
ampl:`model amplAssignment.txt;display x;`	节点 1 的解
ampl:`fix x[1,3]:=0;solve;display x;`	节点 2 的解
ampl:`fix x[2,5]:=0;solve;display x;`	节点 3 的解
ampl:`unfix x[2,5];fix x[5,2]:=0;solve;display x;`	节点 4 的解
ampl:`unfix x[5,2];unfix x[1,3];fix x[3,1]:=0;` `solve;display x;`	节点 5 的解

习题 8.3C

1. 在例 8.3-3 中节点 1 处选择子游程 $2-5-4-2$ 进行分支, 并利用下面两个搜索顺序求解.

(a) 按照优先搜索水平方向的子问题, 在同一层中, 按照从左到右的顺序.

(b) 按照竖直方向进行深度搜索, 直到某个分支已经探明.

***2.** 利用 B&B 算法求解习题 8.3A 中的第 1 题.

3. 利用 B&B 算法求解习题 8.3A 中的第 2 题.

4. 利用 B&B 算法求解习题 8.3A 中的第 3 题.

8.3.3　割平面算法

割平面算法的思想是在指派问题中增加一些约束, 使得求出的解不包含子游程. 附加的约束是按照下面的方法定义的: 对于 n 个城市的例子, 对应于城市 $2,3,\cdots,n$ 引入连续变量 $u_j(\geqslant 0)$, 然后定义下面的附加约束集合为

$$u_i - u_j + nx_{ij} \leqslant n-1, \quad i=2,3,\cdots,n; \quad j=2,3,\cdots,n; \quad i \neq j$$

在指派问题中增加这些约束, 就可以自动避开所有的由子游程组成的解.

例 8.3-5

考虑下面的 4 个城市的 TSP 问题, 距离矩阵是

$$\|d_{ij}\| = \begin{pmatrix} - & 13 & 21 & 26 \\ 10 & - & 29 & 20 \\ 30 & 20 & - & 5 \\ 12 & 30 & 7 & - \end{pmatrix}$$

相应线性规划的约束是由指派问题模型的约束和下表中的附加约束组成, 其中所有的 $x_{ij}=(0,1)$, $u_j \geqslant 0$.

No.	x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}	x_{31}	x_{32}	x_{33}	x_{34}	x_{41}	x_{42}	x_{43}	x_{44}	u_2	u_3	u_4	
1							4										1	−1		$\leqslant 3$
2								4									1		−1	$\leqslant 3$
3										4							−1	1		$\leqslant 3$
4												4						1	−1	$\leqslant 3$
5														4			−1		1	$\leqslant 3$
6															4			−1	1	$\leqslant 3$

最优解是

$$u_2=0, u_3=2, u_4=3, x_{12}=x_{23}=x_{34}=x_{41}=1, \quad \text{游程长度}=59$$

相应的游程是 $1-2-3-4-1$, 很容易证明这个解满足所有附加的约束.

下面我们来看一个子游程解不能满足附加约束的例子, 考虑解 $(1-2-1, 3-4-3)$, 对应的变量 $x_{12}=x_{21}=1, x_{34}=x_{43}=1$. 下面考虑上表中的约束 6:

$$4x_{43}+u_4-u_3 \leqslant 3$$

代入 $x_{43}=1, u_3=2, u_4=3$, 得到 $5 \leqslant 3$, 显然是矛盾的, 所以 $x_{43}=1$ 不成立, 那么也就不能含有子游程 $3-4-3$.

割平面算法的缺点就是随着城市数目的增多, 变量的数目成指数增长, 所以要

得到实际问题的数值解难度很大. 正是这个原因, B&B 算法 (结合启发式算法) 比其他算法更有效一些.

AMPL 程序

图 8.14 给出了割平面算法的 AMPL 模型 (文件 amplEx8.3-5.txt), 模型的数据来自于例 8.3-5 中 4 个城市的 TSP 问题. 程序的输入非常直接: 前两个约束语句表示指派问题模型, 第 3 个语句表示需要增加的割, 用以更换子游程的解. 注意到指派模型的变量都是二元的, 并且在语句 `solve;` 之前要加上 `option solver cplex;` 以保证得到的解是整数.

```
param k;
param n;
param c{1..n,1..n} default 10000;
var x{i in 1..n,j in 1..n} binary;
var u{i in 1..n:i>1}>=0;

minimize tourLength:sum{i in 1..n,j in 1..n}c[i,j]*x[i,j];
subject to
 fromCity {i in 1..n}:sum {j in 1..n} x[i,j] = 1;
 toCity {j in 1..n}:sum {i in 1..n} x[i,j] = 1;
 cut{i in 1..n,j in 1..n:i>1 and j>1 and i<>j}:
                      u[i]-u[j]+n*x[i,j] <= n-1;
data;
param n:=4;
param c:
       1    2    3    4:=
  1    .    13   21   26
  2    10   .    29   20
  3    30   20   .    5
  4    12   30   7    ;
option solver cplex; solve;
display u;
#--------------------------------print formatted output
printf "\n\nOptimal tour length = %7.2f\n",tourLength;
printf "Optimal tour:";
let k:=1;                   #tour starts at city k=1
for {i in 1..n}
    {
    printf "%3i", k;
    for {j in 1..n} #search for next city following k
        {
        if x[k,j]=1 then
           {
           let k:=j; #next city found, set k=j
           break;
           }
        }
    printf "-";              #insert last hyphen
    }
printf " 1\n\n";
```

图 8.14 求解 TSP 问题的 AMPL 割平面模型 (文件 amplEx8.3-5.txt)

程序后半部分的for和if-then语句是为了格式化最后的输出结果：

```
Optimal tour length = 59.00
Optimal tour: 1- 2- 3- 4- 1
```

习题 8.3D

1. 一家工厂使用一种自动导航车给 5 个部门送信件, 导航车从信件收发室出发经过每一个部门, 然后回到收发室. 将收发室的坐标定位为原点 $(0,0)$, 5 个部门的坐标分别是 $(10,30)$, $(10,50)$, $(30,10)$, $(40,40)$, $(50,60)$. 所有坐标点之间的距离单位是米. 自动导航车只能沿着水平或者竖直方向的走廊移动. 目标就是要最小化导航车游程的长度.

   将这个问题建立成 TSP 模型, 并求出需要添加的割.

2. 写出下面 TSP 问题相应的割:

$$\|d_{ij}\| = \begin{pmatrix} \infty & 43 & 21 & 20 & 10 \\ 12 & \infty & 9 & 22 & 30 \\ 20 & 10 & \infty & 5 & 13 \\ 14 & 30 & 42 & \infty & 20 \\ 44 & 7 & 9 & 10 & \infty \end{pmatrix}$$

3. AMPL 实验. 应用 AMPL 的割平面算法求解下面的 TSP 问题.
   (a) 习题 8.3A 中的第 2 题.　　(b) 习题 8.3A 中的第 3 题.

参 考 文 献

Barnett, A., "Misapplication Review: High Roud to Glory", *Interfaces*, Vol. 17, No. 5, pp. 51-54, 1987.

Graves, R., L. Schrage, and J. Sankaran, "An Auction Method for Course Registration", *Interfaces*, Vol. 23, No. 5, pp. 81-97, 1993.

Guéret, C., c. Prins, and M, Sevaux, *Applications of Optimization with Xpress-MP*, Dash Optimization, London, 2002.

Jarvis, J., R. Rardin, V. Unger, R. Moore, and C. Schimpeler, "Optimal Design of Regional Waste-water System: A Fixed Charge Network Flow Model", *Operations Research*, Vol. 26, No. 4, pp. 538-550, 1978.

Lee, J., *A First Course in Combinatorial Optimization*, Cambridge University Press, 2004.

Liberatore, M., and T. Miller, "A Hierarchial Production Planning System", *Interfaces*, Vol. 15, No.4, pp. 1-11, 1985.

Nemhauser, G., and L. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York, 1988.

Salkin, H., and K. Mathur, *Foundations of Integer Programming*, North-Holland, New York, 1989.

Schrijver, A. *Theory of Linear and integer Programming*, Wiley, New York, 1998.

Taha, H., *Integer Programming: Theory, Applications, and Computations*, Academic Press, Orlando, FL, 1975.

Weber, G., "Puzzle Contests in MS/OR Education", *Interfaces*, Vol. 20, No. 2, pp. 72-76, 1990.

Wolsey, L., *Integer Programming*, Wiley, New York, 1998.

第9章 确定性动态规划

本章导读 动态规划 (Dynamic Programming, DP) 通过把一个多变量问题分解成若干个阶段, 每个阶段组成为一个单变量的子问题, 来求出这个多变量问题的最优解. 这种分解的好处是, 每个阶段的优化问题仅涉及单一变量, 比起同时处理多个变量, 计算上更加简单. DP 模型基本上是一种递归方程, 把问题的各阶段联系起来, 保证每个阶段的最优可行解对于整个问题既是最优的也是可行的. 递归方程的表示和概念框架我们前面没有接触过. 经验表明, 对于初学者, 递归方程的结构似乎并不"合乎逻辑". 假如你有类似的体验, 最好的做法是尽量做一些你觉得合乎逻辑的东西, 然后做些适当的计算. 你会很快发现, 本书的定义是正确的, 在这个过程中, 你会了解到 DP 的原理. 我们还为一些例题准备了两个半自动化的 Excel 电子表格, 使用者需要提供关键的信息来启动这些 DP 计算. 这种练习应该能帮助你理解 DP 的一些精微之处.

虽然递归方程是构建 DP 模型的一种通用框架, 但求解的细节各有不同. 只有通过学习不同的 DP 模型, 你才能够获得 DP 建模和求解的经验. 本章给出了一批*确定性的*DP 应用实例, 第 22 章介绍了*随机性的*DP 应用. 在第 10 章和第 15 章中, 还给出了库存模型这一重要领域中的其他应用实例.

本章包括 1 个实际应用实例的概述、7 个解题的例子、2 个 Excel 电子表格模型、32 个节后的习题, 以及 1 个案例. 该案例放在本书 (下册) 的附录 E 中, AMPL/Excel/Solver/TORA 程序放在 ch9Files 文件夹下.

实际应用——Weyerhaeuser 切割及原木分配的优化问题

成材的树木被砍伐下来, 横切成原木, 用来加工成不同的最终产品 (如建筑木材、合板、薄板、纸张等). 原木的规格 (例如长度和两端的直径) 根据使用这些原木的木材场地的要求各有不同. 砍伐的树木高度最高有 100 英尺, 满足木材场地要求的原木横切方案的组合数可能非常大, 而且, 一棵树木的切割方式可能会影响到收益. 目标是确定各种横切的组合方式, 使得总收益最大. 这项研究利用动态规划对这个过程进行优化, 所提出的系统于 1978 年首次实施, 每年增加利润收入达到 700 余万美元. 第 24 章中的案例 8 给出了这项研究的细节.

9.1 DP 计算的递归性质

动态规划的计算是递归进行的, 以便让一个子问题的最优解作为下一个子问题

的输入. 当最后一个子问题求解完成, 也就得到了整个问题的最优解. 递归计算的实现方式, 取决于对原始问题的分解. 尤其是, 这些子问题通常是由共同约束条件联系在一起的. 因为我们要从一个子问题算到下一个子问题, 这些共同约束的可行性必须得到满足.

例 9.1-1 (最短路径问题)

假定你要在两个城市之间选择一条最短的公路线, 图 9.1 给出了从出发城市节点 1 到目的地城市节点 7 的所有可能的路径. 节点 2 到节点 6 表示这些路径通过的中间城市.

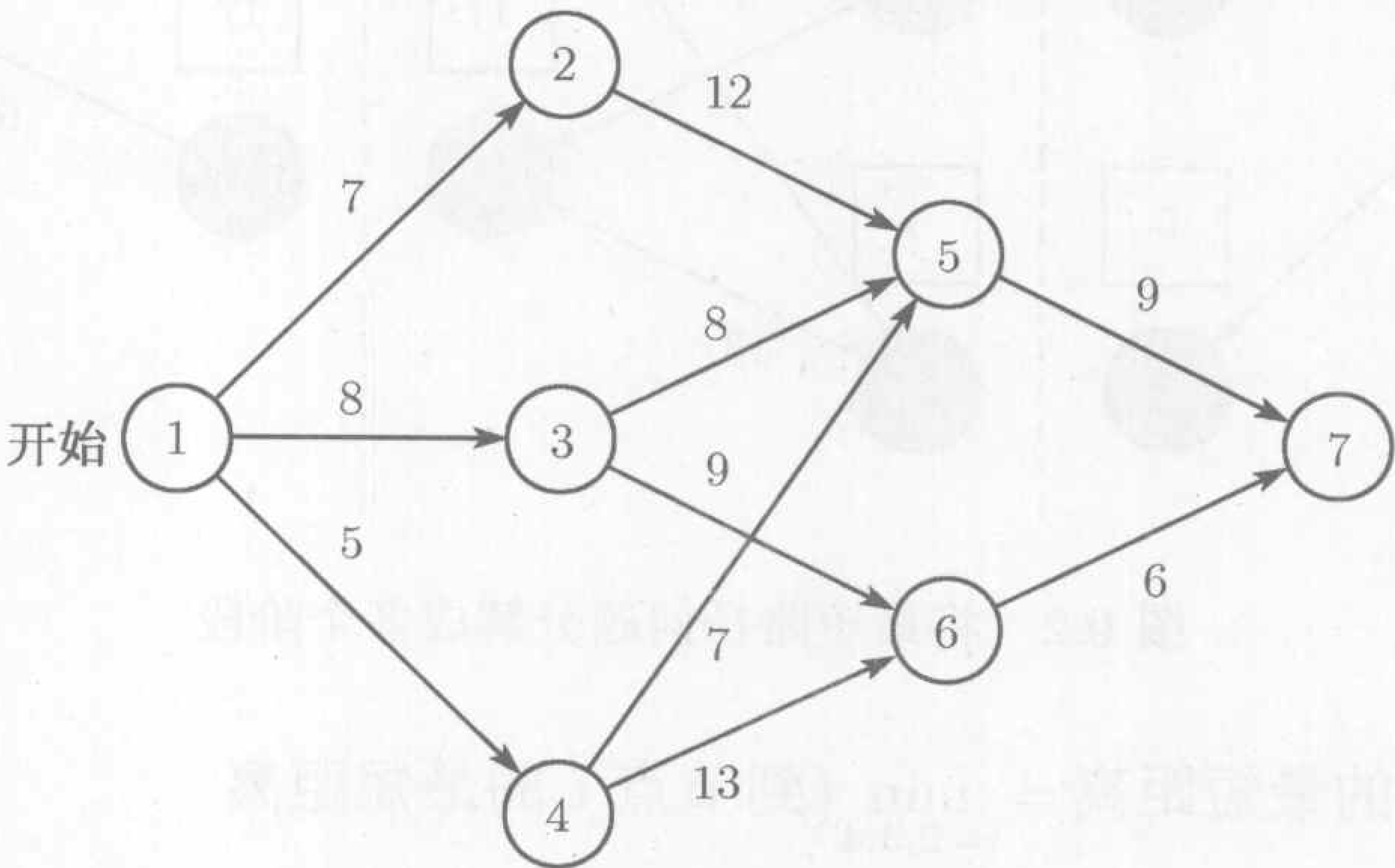


图 9.1 例 9.1-1 的路径网络

可以通过穷举出从节点 1 到节点 7 的所有路径 (一共有 5 条路径) 的办法来解决这个问题, 但对于大的网络, 穷举法的计算量大得难以处理.

为了用 DP 来求解这个问题, 首先用图 9.2 所示的垂直虚线把它分解成若干**阶段**, 然后对每个阶段分别进行计算.

确定最短路径的一般思路是, 对一个阶段的所有终止点, 算出最短的 (累计) 距离, 然后利用这些距离作为输入数据, 输入到下一个后续阶段. 从节点 1 开始, 阶段 1 包括 3 个终点 (2, 3, 4), 它的计算非常简单.

阶段 1 小结

从节点 1 到节点 2 的最短距离 =7 英里 (从节点 1 出发)
从节点 1 到节点 3 的最短距离 =8 英里 (从节点 1 出发)
从节点 1 到节点 4 的最短距离 =5 英里 (从节点 1 出发)

接下来, 阶段 2 有两个终节点: 节点 5 和节点 6. 先考虑节点 5, 从图 9.2 我们看出, 节点 5 可以从 3 个节点 2, 3, 4 经过 3 条不同的路径 (2, 5), (3, 5), (4, 5) 达到, 这些信息, 加上到节点 2, 3, 4 的最短距离, 我们就能确定出到达节点 5 的 (累加) 最短距离为

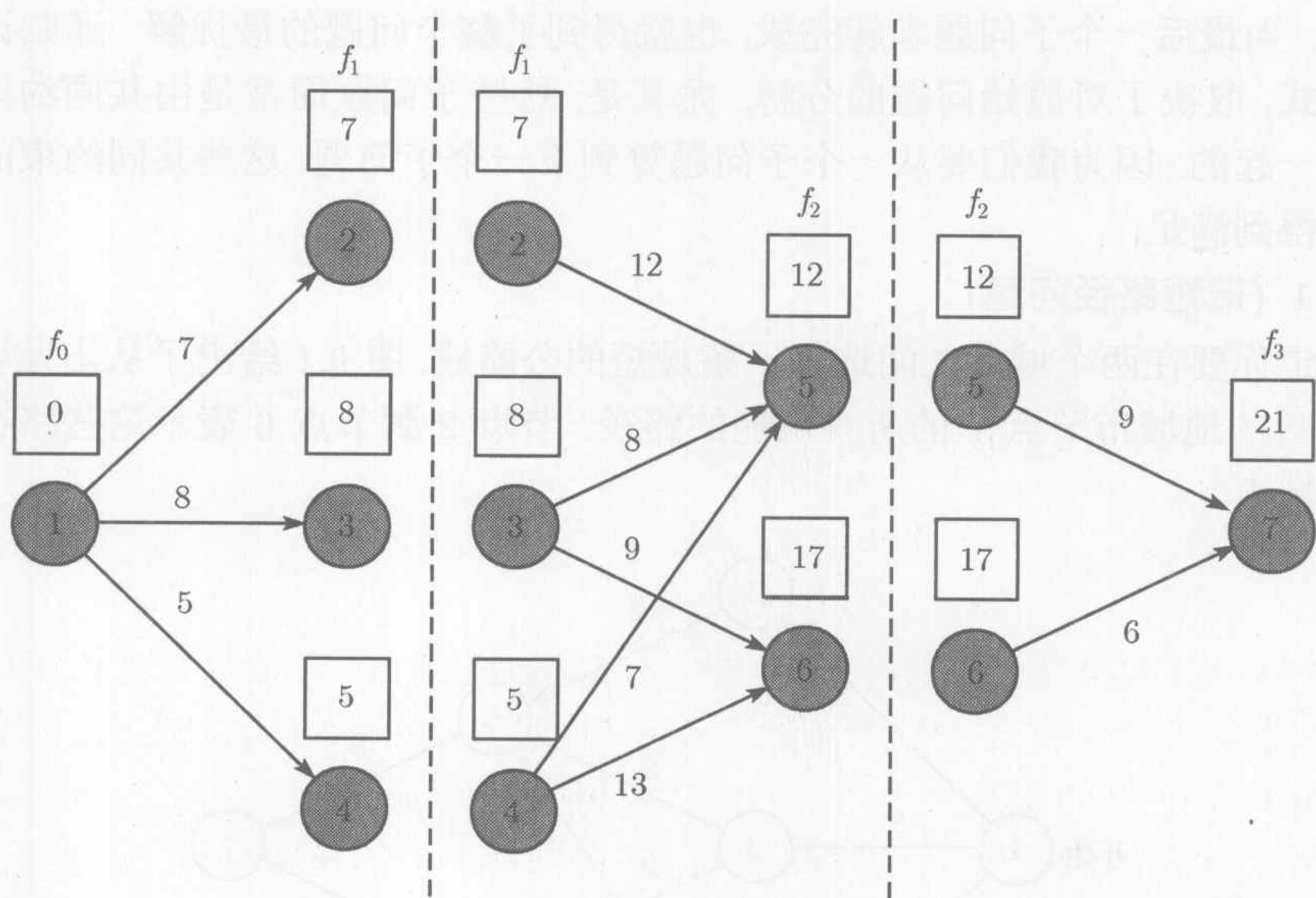


图 9.2　将最短路径问题分解成多个阶段

$$\text{到节点 5 的最短距离} = \min_{i=2,3,4}\{\text{到节点 } i \text{ 的最短距离} + \text{从节点 } i \text{ 到节点 5 的最短距离}\}$$

$$= \min\left\{\begin{array}{c} 7+12=19 \\ 8+8=16 \\ 5+7=12 \end{array}\right\} = 12 \qquad (\text{从节点 4 出发})$$

节点 6 只能从节点 3 和节点 4 到达, 因此

$$\text{到节点 6 的最短距离} = \min_{i=3,4}\{\text{到节点 } i \text{ 的最短距离} + \text{从节点 } i \text{ 到节点 6 的最短距离}\}$$

$$= \min\left\{\begin{array}{c} 8+9=17 \\ 5+13=18 \end{array}\right\} = 17 \qquad (\text{从节点 3 出发})$$

阶段 2 小结

从节点 1 到节点 5 的最短距离 =12 英里 (从节点 4 出发)
从节点 1 到节点 6 的最短距离 =17 英里 (从节点 3 出发)

最后一步来考虑阶段 3. 最终节点 7 可从节点 5 或从节点 6 到达, 利用阶段 2 计算结果和从节点 5 和节点 6 到达节点 7 的距离, 我们得到

$$\text{到节点 7 的最短距离} = \min_{i=5,6}\{\text{到节点 } i \text{ 的最短距离}$$

$$+\text{从节点 } i \text{ 到节点 7 的最短距离}\}$$

$$= \min \left\{ \begin{array}{l} 12 + 9 = 21 \\ 17 + 6 = 23 \end{array} \right\} = 21 \qquad (\text{从节点 5 出发})$$

阶段 3 小结

从节点 1 到节点 7 的最短距离 =21 英里 (从节点 5 出发)

阶段 3 小结表明, 节点 1 与节点 7 之间的最短距离为 21 英里. 为了得到最优的路径, 阶段 3 小结把节点 7 连到节点 5, 阶段 2 小结把节点 4 连到节点 5, 而阶段 1 小结把节点 4 连到节点 1. 这样, 最短的路径是 $1 \to 4 \to 5 \to 7$.

这个例子展现了 DP 中计算的基本特性:

(1) 每个阶段所做的计算都是该阶段可行路径的函数, 并且只针对该阶段.

(2) 当前阶段仅仅连接到*紧接着的*上一阶段, 与再前面的阶段无关. 这种连接是以最短距离小结的形式表示出上一阶段的输出.

递归公式 现在说明如何用数学式子来表示例 9.1-1 中的递归计算. 令 $f_i(x_i)$ 表示阶段 i 到达节点 x_i 的最短距离, 定义 $d(x_{i-1}, x_i)$ 为从节点 x_{i-1} 到节点 x_i 的距离, 则 f_i 是利用下面的递归关系从 f_{i-1} 计算出来的:

$$f_i(x_i) = \min_{\text{所有可行的 } (x_{i-1}, x_i) \text{ 路径}} \{d(x_{i-1}, x_i) + f_{i-1}(x_{i-1})\}, i = 1, 2, 3$$

从 $i = 1$ 开始, 递归设置 $f_0(x_0) = 0$. 这个方程说明, 阶段 i 的最短距离 $f_i(x_i)$ 必须用下一个节点 x_i 来表示. 用动态规划的术语来说, x_i 被看成是系统在阶段 i 的**状态**. 实际上, 阶段 i 的系统*状态*是把这些阶段联系在一起的信息, 使得对余下阶段做出最优决策时, 用不着重新考察前面阶段所做出的决策. 对*状态*做出适当的定义, 能让我们分别地考虑每一个阶段, 并能保证所得出的解对所有阶段都是可行的.

*状态*的这一定义为动态规划产生了后面统一的框架.

最优性原则

对以后阶段所做出的未来决策将会产生一个最优策略, 它与前面各阶段所采用的策略无关.

实现这一原则在例 9.1-1 的计算中是显然的. 例如在阶段 3, 我们只对节点 5 和节点 6 计算最短距离, 而不用管这些节点是怎么从节点 1 过来的. 尽管这一最优性原则对于每个阶段如何优化的细节是"含糊的", 但运用这一原则将为许多复杂问题的求解提供极大的方便.

习题 9.1A

*1. 利用下列路径求解例 9.1-1:

$$d(1,2)=5,\ d(1,3)=9,\ d(1,4)=8$$
$$d(2,5)=10,\ d(2,6)=17$$
$$d(3,5)=4,\ d(3,6)=10$$
$$d(4,5)=9,\ d(4,6)=9$$
$$d(5,7)=8$$
$$d(6,7)=9$$

2. 我是一名户外运动爱好者, 去年夏天, 我和朋友 G.Don 花 5 天时间去了新罕布什尔州风景秀丽的白山 (White Mountains) 山脉登山和露营. 我们当时决定, 到 3 座著名的山峰附近去爬: 华盛顿山、杰弗逊山和亚当山. 华盛顿山从山脚到山顶有 6 英里远, 而杰弗逊山和亚当山相应地分别为 4 英里和 5 英里远. 通向这 3 座山脚下有山路相连, 从华盛顿山到杰弗逊山有 3 英里, 从杰弗逊山到亚当山有 2 英里, 而亚当山到华盛顿山之间有 5 英里远. 我们第 1 天从华盛顿山脚下出发, 第 5 天晚上返回到原地. 我们的目标是尽量走得最远. 我们还商量好了每天只爬一座山, 并在次日开始爬的山的山脚下宿营. 此外, 我们决定任何连着的两天不会去爬同一座山峰. 那么我们是如何安排这次爬山路线的呢?

9.2 前向递归与后向递归

例 9.1-1 采用**前向递归**的方法, 从阶段 1 到阶段 3 向前计算. 这个例子同样可以用**后向递归**的方法来求解, 从阶段 3 开始, 到阶段 1 结束.

前向递归和后向递归都能得出相同的解. 虽然前向递归方法似乎更符合逻辑, 但在 DP 文献中普遍采用后向递归. 原因是, 后向递归一般来说在计算上更快些. 我们下面会用后向递归的方法来演示求解例 9.1-1. 通过这个演示, 你还可以了解到如何用更简洁的表格形式来进行 DP 的计算.

例 9.2-1

例 9.1-1 的后向递归方程为

$$f_i(x_i)=\min_{\text{所有可行的 }(x_i,x_{i+1})\text{ 路径}}\{d(x_i,x_{i+1})+f_{i+1}(x_{i+1})\},i=1,2,3$$

其中当 $x_4=7$ 时 $f_4(x_4)=0$. 相应的计算顺序为 $f_3\to f_2\to f_1$.

阶段 3 由于节点 7($x_4=7$) 分别只有一条路径连接到节点 5 和节点 6($x_3=5$ 和 6), 没有其他选择, 因此阶段 3 的结果可归结为

x_3	$d(x_3,x_4)$	最优解	
	$x_4=7$	$f_3(x_3)$	x_4^*
5	9	9	7
6	6	6	7

阶段 2 路径 (2,6) 由于不存在而断了, 由于阶段 3 给出了 $f_3(x_3)$, 我们可以按照下表对各可行路径进行比较:

x_2	$d(x_2,x_3)+f_3(x_3)$		最优解	
	$x_3=5$	$x_3=6$	$f_2(x_2)$	x_3^*
2	12+9=21	—	21	5
3	8+9=17	9+6=15	15	6
4	7+9=16	13+6=19	16	5

阶段 2 的最优解可解释成: 假如你在城市 2 或城市 4, 最短路径将经过城市 5; 如果你在城市 3, 则最短路径经过城市 6.

阶段 1 从节点 1, 我们有 3 条可选路径: 路径 (1,2), (1,3), (1,4). 利用阶段 2 的 $f_2(x_2)$, 我们可以计算出下表.

x_1	$d(x_1,x_2)+f_2(x_2)$			最优解	
	$x_2=2$	$x_2=3$	$x_2=4$	$f_1(x_1)$	x_2^*
1	7+21=28	8+15=23	5+16=21	21	4

阶段 1 的最优解表示, 城市 1 连到城市 4. 接下来, 阶段 2 的最优解把城市 4 连到城市 5. 最后, 阶段 3 的最优解把城市 5 连到城市 7. 因此最优的整条路径为 $1\to4\to5\to7$, 相应的距离为 21 英里.

习题 9.2A

1. 对习题 9.1A 的第 1 题, 构造出后向递归方程, 并用它来找到最优解.

2. 对习题 9.1A 的第 2 题, 构造出后向递归方程, 并用它来找到最优解.

***3.** 对图 9.3 的网络, 要求计算出城市 1 到城市 7 的最短路径. 利用后向递归定义出各个阶段和状态, 并求解该问题.

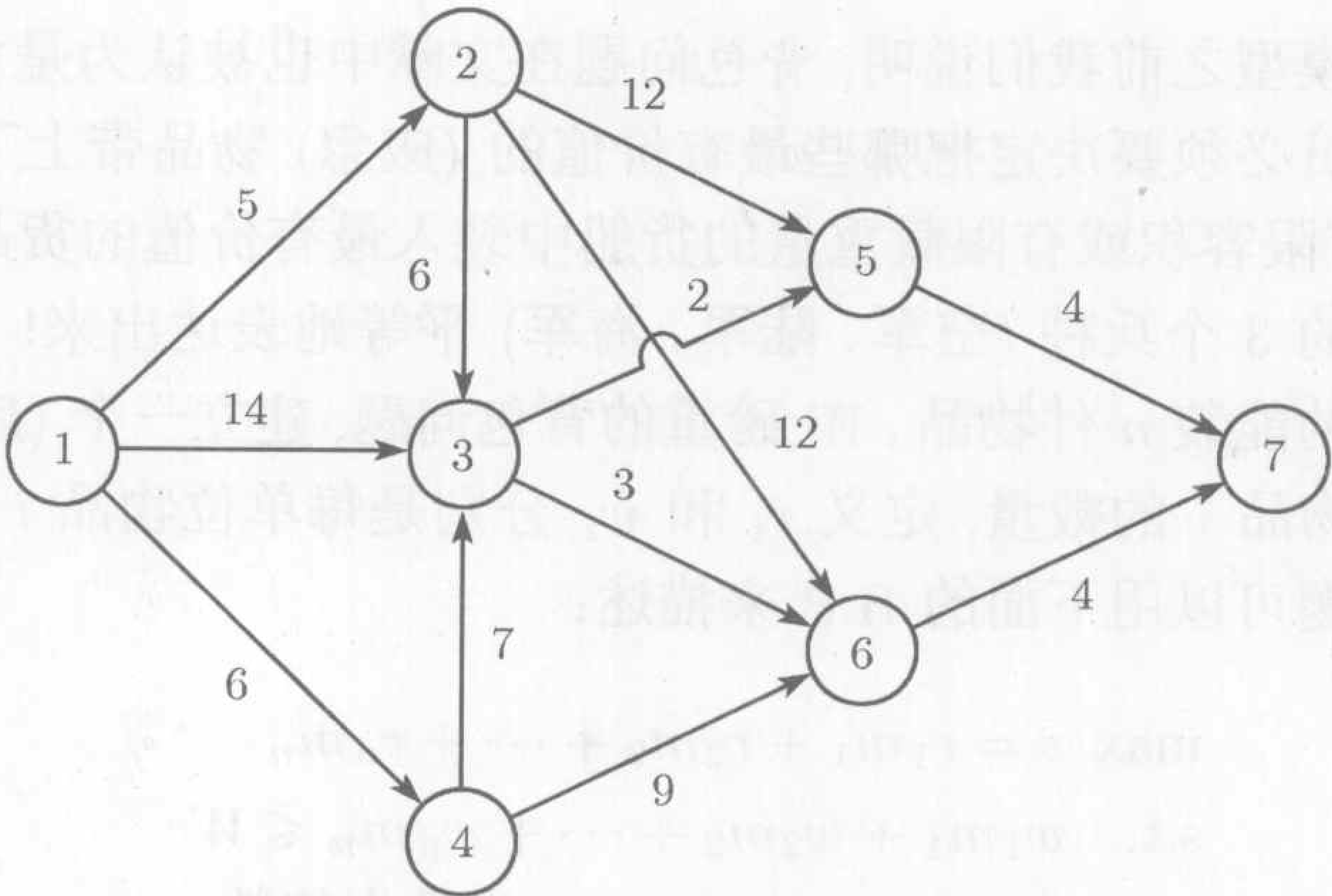


图 9.3 习题 9.2A 第 5 题的网络

9.3　DP 应用选讲

本节介绍 4 个应用, 每项应用都体现了运用动态规划的一种新思路. 在学习每项应用时, 特别要关注 DP 模型的 3 个基本要素:

(1) 定义*阶段*;

(2) 定义每个阶段的*可选方案*;

(3) 定义每个阶段的*状态*.

这 3 个要素中, *状态*的定义往往是最微妙的. 这里介绍的应用表明, 状态的定义根据要建模的实际情况会有很大的不同. 不过, 随着对每项应用问题的深入研究, 你会发现考虑如下问题会大有帮助:

(1) 是什么样的关系把各个阶段联系在一起?

(2) 要在当前的阶段做出可行的决策, 而又不用重新考察前面阶段所做出的决策, 这需要什么样的信息?

根据我的教学经验, 为了加深对*状态*这一概念的理解, 可以通过认真思考书中的定义方式是否合理, 也可以自己给出可能"更加合乎逻辑的"不同定义, 然后把它用在递归计算中. 最后你会发现, 这里所给出的定义提供了解决问题的一种正确的方式. 同时, 所提出的思考过程应该能加强你对状态概念的理解.

9.3.1　背包/飞行箱/装船问题的模型

经典的背包问题是, 一个士兵 (或徒步旅行者) 必须决定在背包里要携带哪些价值最大的物品. 这个问题引申出一般性的资源分配模型, 在这一模型里, 某一种有限的资源分配给多个可选方案 (例如有限的资金分配给多个项目), 目标是使得总收益最大.

在提出 DP 模型之前我们说明, *背包问题*在文献中也被认为是*飞行箱问题*, 即一个喷气机飞行员必须要决定把哪些最有价值的 (应急) 物品带上飞机; 或者是装船问题, 即要在有限容积或有限载重量的货船中装入最有价值的货品. 这 3 个名字似乎是要把军队的 3 个兵种 (空军、陆军、海军) 平等地表达出来!

我们对一般的能装 n 件物品、W 磅重的背包问题, 建立一个 (后向) 递归方程. 令 m_i 为背包中物品 i 的数量, 定义 r_i 和 w_i 分别是每单位物品 i 的收益和重量, 这个一般性的问题可以用下面的 ILP 来描述:

$$
\begin{aligned}
\max\ & z = r_1m_1 + r_2m_2 + \cdots + r_nm_n \\
\text{s.t.}\ & w_1m_1 + w_2m_2 + \cdots + w_nm_n \leqslant W \\
& m_1, m_1, \cdots, m_n \geqslant 0 \text{ 并且为整数}
\end{aligned}
$$

模型的 3 个要素是

(1) 用物品 i 表示阶段 $i, i=1,2,\cdots,n$.

(2) 阶段 i 的可能方案用背包所装入的物品 i 的数量 m_i 来表示, 相关的收益为 $r_i m_i$. 定义 $[\frac{W}{w_i}]$ 为小于等于 $\frac{W}{w_i}$ 的最大整数, 则有 $m_i=0,1,\cdots,[\frac{W}{w_i}]$.

(3) 阶段 i 的状态用 x_i 表示, 其含义为放置到阶段 (物品)$i,i+1,\cdots,n$ 的总重量. 这个定义反映出的事实是, 重量约束是把所有 n 个阶段联系在一起的唯一限制条件.

定义

$$f_i(x_i)=\text{给定状态 } x_i \text{ 下, 阶段 } i,i+1,\cdots,n \text{ 的最大收益}$$

确定其递归方程最简单的方式就是下面的两个步骤的算法:

步骤 1 把 $f_i(x_i)$ 表示成下面 $f_{i+1}(x_{i+1})$ 的函数:

$$f_i(x_i)=\min_{\substack{m_i=0,1,\cdots,[\frac{W}{w_i}]\\ x_i\leqslant W}}\{r_i m_i+f_{i+1}(x_{i+1})\},\quad i=1,2,\cdots,n$$

$$f_{n+1}(x_{n+1})\equiv 0$$

步骤 2 将 x_{i+1} 表示成 x_i 的函数, 以保证左边项 $f_i(x_i)$ 仅是 x_i 的函数. 按照定义, $x_i-x_{i+1}=w_i m_i$ 表示阶段 i 所用的重量, 因此 $x_{i+1}=x_i-w_i m_i$, 合理的递归方程为

$$f_i(x_i)=\max_{\substack{m_i=0,1,\cdots,[\frac{W}{w_i}]\\ x_i\leqslant W}}\{r_i m_i+f_{i+1}(x_i-w_i m_i)\},\quad i=1,2,\cdots,n$$

例 9.3-1

一艘载货量 4 吨的货船可装载 3 种货物. 下表给出每种货物 i 的单位重量 w_i(单位是吨) 以及单位收益 r_i(单位是 1 000 美元). 该货船应如何装载这些货物才能使得收入最大呢?

货物 i	w_i	r_i
1	2	31
2	3	47
3	1	14

因为单位重量 w_i 和最大重量 W 为整数, 状态 x_i 也一定只取整数值.

阶段 3 阶段 3(货物 3) 分配的具体重量是未知的, 但一定取值 $0,1,\cdots,4$ 之一 (因为 $W=4$ 吨). 状态 $x_3=0$ 和 $x_3=4$ 分别表示根本没有运送货物 3 和整船全部装载货物 3 这两种极端情况. 其余的 $x_3(=1,2,3)$ 的值意味着货船载货容量有一部分分配给了货物 3. 实际上, x_3 的取值范围覆盖了货船装载货物 3 的所有可能情况.

已知 $w_3=1$ 吨/单位, 可装载货物 3 的最大单位数为 $\frac{4}{1}=4$, 这意味着 m_3 的可能取值为 $0,1,2,3,4$. m_3 是可行的, 仅当 $w_3 m_3\leqslant x_3$. 因此, 所有不可行的装载方

案 (对应于 $w_3m_3 > x_3$) 都被排除了. 以下方程是比较阶段 3 各种方案的依据.

$$f_3(x_3) = \max_{m_3=0,1,\cdots,4} \{14m_3\}$$

下表比较 x_3 每个不同取值的可行方案.

x_3	$14m_3$					最优解	
	$m_3=0$	$m_3=1$	$m_3=2$	$m_3=3$	$m_3=4$	$f_3(x_3)$	m_3^*
0	0	—	—	—	—	0	0
1	0	14	—	—	—	14	1
2	0	14	28	—	—	28	2
3	0	14	28	42	—	42	3
4	0	14	28	42	56	56	4

阶段 2　$\max\{m_2\} = [\frac{4}{3}] = 1$, 或 $m_2 = 0, 1$

$$f_2(x_2) = \max_{m_2=0,1} \{47m_2 + f_3(x_2 - 3m_2)\}$$

x_2	$47m_2 + f_3(x_2 - 3m_2)$		最优解	
	$m_2=0$	$m_2=1$	$f_2(x_2)$	m_2^*
0	$0+0=0$	—	0	0
1	$0+14=14$	—	14	0
2	$0+28=28$	—	28	0
3	$0+42=42$	$47+0=47$	47	1
4	$0+56=56$	$47+14=61$	61	1

阶段 1　$\max\{m_1\} = [\frac{4}{2}] = 2$, 或 $m_1 = 0, 1, 2$

$$f_1(x_1) = \max_{m_1=0,1,2} \{31m_1 + f_2(x_1 - 2m_1)\}, \max\{m_1\} = \left[\frac{4}{2}\right] = 2$$

x_1	$31m_1 + f_2(x_1 - 2m_1)$			最优解	
	$m_1=0$	$m_1=1$	$m_1=2$	$f_1(x_1)$	m_1^*
0	$0+0=0$	—	—	0	0
1	$0+14=14$	—	—	14	0
2	$0+28=28$	$31+0=31$	—	31	1
3	$0+47=47$	$31+14=45$	—	47	0
4	$0+61=61$	$31+28=59$	$62+0=62$	62	2

最优解按照如下方法得到：由于 $W = 4$ 吨, 根据阶段 1, $x_1 = 4$ 给出 $m_1^* = 2$ 的最优分配方案, 表示船上将要装入 2 个单位的物品 1. 这一方案使得 $x_2 = x_1 - 2m_2^* = 4 - 2\times 2 = 0$. 从阶段 2 得知, 如果 $x_2 = 0$, 则 $m_2^* = 0$, 这又使得 $x_3 = x_2 - 3m_2 = 0 - 3\times 0 = 0$. 接下来, 根据阶段 3, $x_3 = 0$ 给出 $m_3^* = 0$. 因此, 整个最优解为 $m_1^* = 2, m_2^* = 0, m_3^* = 0$, 相应收益为 $f_1(4) = \$62\,000$.

从阶段 1 的表中, 我们实际上需要得出 $x_1 = 4$ 的最优解, 因为这是最后一个要考虑的阶段. 但是, 对于 $x_1 = 0, 1, 2, 3$ 的计算也包括进来了, 以便进行*灵敏度分析*. 例如, 假如货船载货量是 3 吨而不是 4 吨时会怎么样? 在这种情况下会得出下面一个新的最优解:

$$(x_1 = 3) \to (m_1^* = 0) \to (x_2 = 3) \to (m_2^* = 1) \to (x_3 = 0) \to (m_3^* = 0)$$

这种情况下, 最优解向量为 $(m_1^*, m_2^*, m_3^*) = (0, 1, 0)$, 最优收益 $f_1(3) = \$47\,000$.

评注 这个装船问题的例子代表了一类典型的*资源分配模型*, 这类模型把某种有限的资源在有限数目的 (经济) 活动之间进行分配. 分配的目标是使得某个相应的收入函数达到极大值. 在这类模型中, 每个阶段状态的定义都类似于我们装船模型给出的定义, 即阶段 i 的状态就是分配到阶段 $i, i+1, \cdots, n$ 的资源总量.

Excel 程序

由于动态规划计算的特殊性质, 不可能有一个通用计算机程序来处理所有的 DP 问题. 大概正是因为这一点, 一直没有开发出商业化的动态规划软件包.

本节介绍一种利用 Excel 开发的算法, 来处理专门一类 DP 问题: 单一约束的背包问题 (文件 Knapsack.xls). 该算法可针对各种数据, 能处理最多有 10 种装包物品的问题.

图 9.4 是这一背包问题的 (后向递归) 动态规划模型的开始界面, 它分成两部分: 右边部分 (Q:V) 用来汇总输出的解, 左边部分 (A:P) 的第 3, 4, 6 行为当前阶段提供输入数据, 第 7 行以下留作阶段的计算. 这些输入数据的符号采用了 DP 模型中的数学符号, 一看就可以知道其含义. 为了方便在一屏上显示这个电子表格, 阶段 i 的物品数 m_i 的最大可行值设定为 10(单元格 D6:N6).

	A	B	C	D	E	F	G	H	O	P	Q	R	S	T	U	V
1	Dynamic Programming (Backward) Knapsack Model															
2	Input Data and Stage Calculations										Ouput Solution Summary					
3	Number of stages,N=				Res. limit, W=						x	f	m	x	f	m
4	Current stage=			w=	1	r=										
5	Are m values correct?								Stage Optimum Solution							
6			m=													
7			r*m=													
8			w*m=						f	m						
9																
10																
11																

图 9.4 一般 DP 背包问题模型的 Excel 开始界面 (文件 excelKnapsack.xls)

图 9.5 显示了该算法对例 9.3-1 生成的阶段计算结果. 这些计算每次针对一个阶段进行, 由用户提供每个阶段计算的基本数据. 不妨利用这种方式来试一试, 它可以增加你对 DP 计算详细过程的理解.

Stage 3 :

	A	B	C	D	E	F	G	H	O	P	Q	R	S	T	U	V
1	Dynamic Programming (Backward) Knapsack Model															
2	Input Data and Stage Calculations										Ouput Solution Summary					
3	Number of stages,N=			3	Res. limit, W=		4				x	f	m	x	f	m
4	Current stage=		3	w3=	1	r3=	14				stage 3					
5	Are m3 values correct?			yes	yes	yes	yes	yes	Stage		0	0	0			
6			m3=	0	1	2	3	4	Optimum		1	14	1			
7	Stage4		r3*m3=	0	14	28	42	56	Solution		2	28	2			
8	f4		w3*m3=	0	1	2	3	4	f3	m3	3	42	3			
9		x3=	0	0	111111	111111	111111	111111	0	0	4	56	4			
10		x3=	1	0	14	111111	111111	111111	14	1						
11		x3=	2	0	14	28	111111	111111	28	2						
12		x3=	3	0	14	28	42	111111	42	3						
13		x3=	4	0	14	28	42	56	56	4						

Stage 2 :

	A	B	C	D	E	F	G	H	O	P	Q	R	S	T	U	V
1	Dynamic Programming (Backward) Knapsack Model															
2	Input Data and Stage Calculations										Ouput Solution Summary					
3	Number of stages,N=			3	Res. limit, W=		4				x	f	m	x	f	m
4	Current stage=		2	w2=	3	r2=	47				stage 3			stage 2		
5	Are m2 values correct?			yes	yes	delete	delete	delete	Stage		0	0	0	0	0	0
6			m2=	0	1	2	3	4	Optimum		1	14	1	1	14	0
7	Stage3		r2*m2=	0	47				Solution		2	28	2	2	28	0
8	f3		w2*m2=	0	3				f2	m2	3	42	3	3	47	1
9	0	x2=	0	0	111111				0	0	4	56	4	4	61	1
10	14	x2=	1	14	111111				14	0						
11	28	x2=	2	28	111111				28	0						
12	42	x2=	3	42	47				47	1						
13	56	x2=	4	56	61				61	1						

Stage 1 :

	A	B	C	D	E	F	G	H	O	P	Q	R	S	T	U	V
1	Dynamic Programming (Backward) Knapsack Model															
2	Input Data and Stage Calculations										Ouput Solution Summary					
3	Number of stages,N=			3	Res. limit, W=		4				x	f	m	x	f	m
4	Current stage-		1	w1-	2	r1-	31				stage 3			stage 2		
5	Are m1 values correct?			yes	yes	yes	delete	delete	Stage		0	0	0	0	0	0
6			m1=	0	1	2	3	4	Optimum		1	14	1	1	14	0
7	Stage2		r1*m1=	0	31	62			Solution		2	28	2	2	28	0
8	f2		w1*m1=	0	2	4			f1	m1	3	42	3	3	47	1
9	0	x1=	0	0	111111	111111			0	0	4	56	4	4	61	1
10	14	x1=	1	14	111111	111111			14	0				stage 1		
11	28	x1=	2	28	31	111111			31	1				0	0	0
12	47	x1=	3	47	45	111111			47	0				1	14	0
13	61	x1=	4	61	69	62			62	2				2	31	1
14														3	47	0
15														4	62	2

图 9.5 解例 9.3-1 背包问题的 Excel 动态规划模型 (文件 excelKnapsack.xls)

从阶段 3 开始, 用例 9.3-1 的符号和数据, 各输入单元格中的值改成如下列表:

单元格	输入值
D3	阶段总数, $N=3$
G3	资源限制, $W=4$
C4	当前阶段 $=3$
E4	$w_3=1$
G4	$r_3=14$
D6:H6	$m_3=(0,1,2,3,4)$

注意到在例 9.3-1 中, m_3 的可行值为 $0,1,\cdots,[\frac{W}{w_3}]=[\frac{4}{1}]=4$, 电子表格会自动告诉你需要多少个 m_3 的值, 并检查你输入的数值是否正确, 在第 5 行显示自检查结果: "yes"、"no" 和 "delete".

在输入了阶段 3 的数据并通过检查以后, 电子表格就会"激活"并 (从 B 列到 P 列) 自动产生该阶段所有必要的计算结果, 值 $-1\,111\,111$ 用来指示对应的输入不可行. O 列和 P 列给出这个阶段的最优解 (f_3, m_3). A 列给出 f_4 的值. 因为计算开始于阶段 3, 所以对于 x_3 的所有取值, $f_4=0$, 因此你可以让 A9:A13 全部空白或全输入零值.

有了阶段 3 的计算结果以后, 按照下列步骤对当前阶段的最优解建立一个永久性纪录, 准备用电子表格进行下一阶段的计算:

第 1 步 复制 C9:C13 中 x_3 的各值, 并粘贴到最优解计算结果部分的 Q5:Q9 各单元格. 接下来, 复制 O9:P3 的 (f_3, m_3) 各值, 粘贴到 R5:S9. 记住, 你只需要粘贴这些数值, 操作方法是在 Edit 菜单下选定 Paste Special, 并在对话框中选择 Values.

第 2 步 复制 R5:R9 中 f_3 的值, 粘贴到 A9:A13 中 (不必用 Paste Special).

第 3 步 把单元 C4 改成 2, 输入 w_2, r_2, m_2 的新值, 记录阶段 2 的数据.

第 2 步是把 $f_{i+1}(x_i - w_i m_i)$ 放到 A 列里面, 准备计算阶段 i 的 $f_i(x_i)$ (见例 9.3-1 背包问题的递归公式). 这就说明了为什么要在阶段 3 的表中用 0 值代表 f_4 的原因.

有了阶段 2 的计算结果后, 可以按照类似的方法准备阶段 1 的界面. 我们在例 9.3-1 中已经解释过, 当阶段 1 算完以后, 就可以从最优解汇总中知道最优解了. 需要注意的是, 对输出解汇总区域 (屏幕右边的 Q:V 列) 的组织是可以自由安排的, 可以根据自己习惯的任何方式来组织这些内容.

习题 9.3A [①]

1. 在例 9.3-1 中, 假设货船的最大载重量分别为 2 吨和 5 吨, 求出最优解.
2. 用下列两组数据, 求出例 9.3-1 的装船问题的最优解:
   *(a) $w_1=4, r_1=70, w_2=1, r_2=20, w_3=2, r_3=40, W=6$
   (b) $w_1=1, r_1=30, w_2=2, r_2=60, w_3=3, r_3=80, W=4$
3. 在例 9.3-1 的装船问题中, 假设每单件货物的收益为某一常数, 而且仅当选中该货品后为已知, 如下表:

---

① 在这组习题中, 建议你尽量手算出一个结果, 然后利用模板 excelKnapsack.xls 对你的结果进行检验.

货品	收益
1	$\begin{cases} -5+31m_1, & \text{如果 } m_1>0 \\ 0, & \text{如果 } m_1 \leqslant 0 \end{cases}$
2	$\begin{cases} -15+47m_2, & \text{如果 } m_2>0 \\ 0, & \text{如果 } m_2 \leqslant 0 \end{cases}$
3	$\begin{cases} -4+14m_3, & \text{如果 } m_3>0 \\ 0, & \text{如果 } m_3 \leqslant 0 \end{cases}$

用动态规划求出最优解. (提示：可以用 Excel 文件 excelSetupKnapsack.xls 来检查你的计算过程.)

4. 一位野外徒步旅行者必须往包里装 3 样东西：食品、急救包和衣物. 背包容积为 3 ft^3 (立方英尺)，每单位食品占空间 1 ft^3，一个急救包占 1/4 ft^3，每件衣物占 1/2 ft^3. 该旅行者对食品、急救包和衣物给出了优先级权重，分别是 3, 4, 5, 意味着衣物在这 3 种东西里价值最高. 根据经验, 他每样东西至少要拿 1 件, 急救包最多需要 2 个. 那么这个徒步旅行者把每样东西到底应该拿多少呢？

***5.** 某学生必须在 4 个不同系里选修 10 门课程, 每个系至少选一门. 选择分布在 4 个系里的 10 门课程, 要争取使获得的“知识量”达到最大. 这个学生按照 100 分制来衡量能学到的知识, 得出了下表：

系号	选课数						
	1	2	3	4	5	6	≥7
I	25	50	60	80	100	100	100
II	20	70	90	100	100	100	100
III	40	60	80	100	100	100	100
IV	10	20	30	40	50	60	70

这位学生应该如何选课呢？

6. 我的后院有一个 10×20 英尺的小花园, 今年春天, 我打算种上 3 种蔬菜：西红柿、豆角和玉米. 花园要安排成 10 排垄, 每条垄 10 英尺长, 要求玉米和西红柿垄宽 2 英尺, 豆角垄宽 3 英尺. 我最爱吃西红柿, 最不爱吃豆角, 按照 10 分制, 我给西红柿打 10 分, 玉米 7 分, 豆角 3 分. 不管我爱不爱吃, 我妻子坚持至少种 1 垄豆角, 最多种 2 垄西红柿. 在这种情况下, 每种蔬菜该种多少排呢？

***7.** 一家非常好的慈善机构通过义工为需要帮助的家庭盖房子. 某个符合条件的家庭可选择 3 种房屋面积标准：1 000, 1 100, 1 200 ft^2，每种规格大小的房屋需要的义工数量不同. 该组织在 Fayetteville 的分部接收到未来半年的 5 份申请, 委员会按照若干因素对每份申请打了分, 越需要的, 分数就越高. 在以后的 6 个月里, Fayetteville 分部可招募到共 23 名义工. 下面的数据是对申请者的打分和所需义工数的情况, 委员会应该批准哪些申请呢？

申请	房屋规格 (ft^2)	分数	所需义工数
1	1 200	78	7
2	1 000	64	4
3	1 100	68	6
4	1 000	62	5
5	1 200	85	8

8. 沙里夫 · 巴萨姆想参加华盛顿县的县长的改选, 筹到了 \$10 000 竞选资金. 虽然改选委员会想在县里的 5 个选区开展竞选活动, 但竞选资金不足. 下表列出了选民数和在每个选区开展有效竞选活动所需的资金额. 对每个选区的选择是, 要么足额分配资金, 要么放弃. 这些竞选资金该如何分配呢?

选区	选民人数	所需资金 (\$)
1	3 100	3 500
2	2 600	2 500
3	3 500	4 000
4	2 800	3 000
5	2 400	2 000

9. 一台电器由 3 个部件组成. 这 3 个部件串联, 假如有一个部件发生故障, 电器就不能工作. 我们可以通过在每个部件里安装 1 到 2 个备份元件来提高该电器的可靠性 (不发生故障的概率). 下表列出了可靠性 r 和成本费用 c. 假设制造该电器的可有资金共 \$10 000, 那么怎么来制造这台电器呢?(**提示**: 目标是使整个电器的可靠性 $r_1r_2r_3$ 达到最大, 这意味着目标函数的分解是乘法的, 而不是加法的.)

并联元件数	部件 1		部件 2		部件 3	
	r_1	c_1(\$)	r_2	c_2(\$)	r_3	c_3(\$)
1	0.6	1 000	0.7	3 000	0.5	2 000
2	0.8	2 000	0.8	5 000	0.7	4 000
3	0.9	3 000	0.9	6 000	0.9	5 000

10. 用动态规划求解以下模型:

$$\max \quad z = \prod_{i=1}^{n} y_i$$

$$\text{s.t.} \quad y_1 + y_2 + \cdots + y_n = c$$

$$y_j \geqslant 0, \quad j = 1, 2, \cdots, n$$

(**提示**: 本题除了变量 y_j 是连续的之外, 和第 9 题类似.)

11. 用 DP 求解以下问题:

$$\min \quad z = y_1^2 + y_2^2 + \cdots + y_n^2$$

$$\text{s.t.} \quad \prod_{i=1}^{n} y_i = c$$

$$y_i > 0, \quad i = 1, 2, \cdots, n$$

12. 用 DP 求解以下问题:

$$\begin{aligned} \max\quad & z=(y_1+2)^2+y_2y_3+(y_4-5)^2 \\ \text{s.t.}\quad & y_1+y_2+y_3+y_4\leqslant 5 \\ & y_i\geqslant 0 \text{ 并且是整数},\quad i=1,2,3,4 \end{aligned}$$

13. 用 DP 解以下问题:

$$\begin{aligned} \min\quad & z=\max\{f(y_1),f(y_2),\cdots,f(y_n)\} \\ \text{s.t.}\quad & y_1+y_2+\cdots+y_n=c \\ & y_i\geqslant 0,\ i=1,2,\cdots,n \end{aligned}$$

对于 $n=3, c=10$, 而且 $f(y_1)=y_1+5, f(y_2)=5y_2+3, f(y_3)=y_3-2$ 的特殊情况, 给出最优解.

9.3.2 劳动力规模模型

在一些建筑工程中, 常用招工和解聘的做法来保证工人的数量满足工程的需要. 招聘和解聘工人肯定都需要额外的费用, 那么如何在工程期间保持足够的劳动力呢?

假设工程工期长度为 n 个星期, 并且第 i 周最少需要 b_i 个劳动力. 理论上讲, 我们可以通过招聘和解聘让第 i 周的劳动力总数恰好等于 b_i. 但是, 一个可能更经济的做法是, 维持某个劳动力规模, 让它大于通过招工方式达到的最小需求数. 这就是我们下面将要考虑的问题.

假定 x_i 为第 i 周雇用的实际工人数, 第 i 周可能出现 2 笔费用: $C_1(x_i-b_i)$ 是维持多余劳动力 (x_i-b_i) 的费用, $C_2(x_i-x_{i-1})$ 是招聘更多劳动力 (x_i-x_{i-1}) 的费用. 我们还假定如果停止招工的话, 则不需要其他的费用.

下面定义动态规划模型的要素:

(1) 用周 i, $i=1,2,\cdots,n$ 代表阶段i.

(2) 阶段 i 的*备选方案*为 x_i, 它是第 i 周的工人数.

(3) 阶段 i 的*状态*用第 $(i-1)$ 阶段 (周) 已有的劳动力数 x_{i-1} 来表示.

这个 DP 模型的递归关系如下:

$$f_i(x_{i-1})=\min_{x_i\geqslant b_i}\{C_1(x_i-b_i)+C_2(x_i-x_{i-1})+f_{i+1}(x_i)\}, i=1,2,\cdots,n$$

$$f_{n+1}(x_n)\equiv 0$$

从阶段 n 开始, 用 $x_n=b_n$ 开始计算, 到阶段 1 停止.

例 9.3-2

一建筑工程承包商估计, 在未来的 5 周内, 所需的劳动力数量分别为 5, 7, 8, 4

和 6 个工人. 对多出的工人, 每人每周的成本费用是 \$300. 不论在哪一周, 新雇用一名工人的固定费用是 \$400, 外加每人每周 \$200.

这个问题的数据汇总如下:

$$b_1=5,\ b_2=7,\ b_3=8,\ b_4=4,\ b_5=6$$
$$C_1(x_i-b_i)=3(x_i-b_i),\ x_i>b_i,\quad i=1,2,\cdots,5$$
$$C_2(x_i-x_{i-1})=4+2(x_i-x_{i-1}),\ x_i>x_{i-1},\quad i=1,2,\cdots,5$$

费用函数 C_1 和 C_2 的单位为 100 美元.

阶段 5 $(b_5=6)$

x_4	$C_1(x_5-6)+C_2(x_5-x_4)$	最优解	
	$x_5=6$	$f_5(x_4)$	x_5^*
4	$3(0)+4+2(2)=8$	8	6
5	$3(0)+4+2(1)=6$	6	6
6	$3(0)+0 \quad =0$	0	6

阶段 4 $(b_4=4)$

x_3	$C_1(x_4-4)+C_2(x_4-x_3)+f_5(x_4)$			最优解	
	$x_4=4$	$x_4=5$	$x_4=6$	$f_4(x_3)$	x_4^*
8	$3(0)+0+8=8$	$3(1)+0+6=9$	$3(2)+0+0=6$	6	6

阶段 3 $(b_3=8)$

x_2	$C_1(x_3-8)+C_2(x_3-x_2)+f_4(x_3)$	最优解	
	$x_3=8$	$f_3(x_2)$	x_3^*
7	$3(0)+4+2(1)+6=12$	12	8
8	$3(0)+0 \quad +6=6$	6	8

阶段 2 $(b_2=7)$

x_1	$C_1(x_2-7)+C_2(x_3-x_2)+f_3(x_2)$		最优解	
	$x_2=7$	$x_2=8$	$f_2(x_1)$	x_2^*
5	$3(0)+4+2(2)+12=20$	$3(1)+4+2(3)+6=19$	19	8
6	$3(0)+4+2(1)+12=18$	$3(1)+4+2(2)+6=17$	17	8
7	$3(0)+0 \quad +12=12$	$3(1)+4+2(1)+6=12$	12	7
8	$3(0)+0 \quad +12=12$	$3(1)+0 \quad +6=9$	9	8

阶段 1 $(b_1=5)$

x_0	$C_1(x_1-5)+C_2(x_1-x_0)+f_2(x_1)$				最优解	
	$x_1=5$	$x_1=6$	$x_1=7$	$x_1=8$	$f_1(x_0)$	x_1^*
0	$3(0)+4+2(5)$ $+19=33$	$3(1)+4+2(6)$ $+17=36$	$3(2)+4+2(7)$ $+12=36$	$3(2)+4+2(8)$ $+9=35$	33	5

得到的最优解为

$$x_0 = 0 \to x_1^* = 5 \to x_2^* = 8 \to x_3^* = 8 \to x_4^* = 6 \to x_5^* = 6$$

这个解可以解释为下列计划:

第 i 周	最少工人数 (b_i)	实际工人数 (x_i)	决策方案	费用
1	5	5	新雇 5 名工人	$4+2\times5=14$
2	7	8	新雇 3 名工人	$4+2\times3+1\times3=13$
3	8	8	不变	0
4	4	6	解雇 2 名工人	$3\times2=6$
5	6	6	不变	0

总费用为 $f_1(0)=\$3\ 300$.

习题 9.3B

1. 针对下列每一组最少劳动力需求, 求解例 9.3-2:

*(a) $b_1=6,\ b_2=5,\ b_3=3,\ b_4=6,\ b_5=8$

(b) $b_1=8,\ b_2=4,\ b_3=7,\ b_4=8,\ b_5=2$

2. 例 9.3-2 中, 如果对每位解雇的工人要支付补偿费 \$100, 请求出最优解.

***3.** Luxor 旅行社安排埃及南部 7 日游线路. 按协议, 旅行社将在未来 4 个星期里为旅行团提供租车, 数量分别为 7 辆、4 辆、7 辆和 8 辆. Luxor 旅行社和当地一家车行签了合同, 为游客提供这些车辆. 车行每周每车收租金 \$220, 外加收取一次性租车交易费 \$500. 但是, Luxor 旅行社可以决定周末不还车, 这样只要付每周的租金 (\$220). Luxor 旅行社怎么才能以最好的方式处理这个租车问题呢?

4. GECO 公司按照合同在未来 4 年内每年生产 4 台飞机发动机, 每年的生产能力和生产成本都不同. GECO 在第 1 年能生产 5 台发动机, 第 2 年生产 6 台, 第 3 年生产 3 台, 第 4 年生产 5 台. 相应地, 这 4 年里每台发动机的生产成本分别是 \$300 000、\$330 000、\$350 000 和 \$420 000. 在每一年里, GECO 也可以决定多生产发动机, 在这种情况下, 发动机必须先存放在库房里, 直到运走. 每年每台发动机的库存成本也不一样, 按照估计, 第 1 年为 \$20 000, 第 2 年为 \$30 000, 第 3 年 \$40 000, 第 4 年是 \$50 000. 现在正处于第一年开头, GECO 有 1 台发动机可以交货. 请为 GECO 制定一个最优的生产计划.

9.3.3　设备更新模型

一台机器用得越久, 维修成本就越高, 生产能力就越低. 当一台机器到了某个年限后, 对其及时更新可能会更加经济. 因此这就成了一个确定最经济的机器更新年限的问题.

假设我们正在研究一个 n 年期间的设备更新问题. 在每年的*年初*, 我们来决定这台机器是要再使用一年, 还是要用一台新的机器来更新它. 令 $r(t)$、$c(t)$ 和 $s(t)$ 分别表示某一台运行了 t 年的机器的年收入、运行费用和折旧现值, 购买一台新机器的费用每年都是 I.

这个 DP 模型的要素是

(1) 用年度 $i, i=1,2,\cdots,n$ 代表阶段 i.

(2) 阶段 (年度)i 的备选方案为在年度 i 年初是继续使用这台机器, 还是进行更新.

(3) 阶段 i 的状态是该台机器在第 i 年开始时已经使用的年数.

假定这台机器在 i 年初的设备运行年数为 t, 定义

$$f_i(t) = \text{年度 } i, i+1, \cdots, n \text{ 的最大净收入}$$

得出的递归关系为

$$f_i(t) = \max \begin{cases} r(t) - c(t) + f_{i+1}(t+1), & \text{如果继续使用} \\ r(0) + s(t) - I - c(0) + f_{i+1}(1), & \text{如果更新} \end{cases}$$

$$f_{n+1}(\cdot) \equiv 0$$

例 9.3-3

某公司需要对一台已经使用了 3 年的机器确定今后 4 年 ($n=4$) 的最优更新策略. 公司要求, 用了 6 年的机器就要更新, 购买一台新机器的价格是 \$100 000. 以下是该问题的数据表 (单位: 美元):

使用年数 t(年)	收入 $r(t)$	运行成本 $c(t)$	折旧现值 $s(t)$
0	20 000	200	—
1	19 000	600	80 000
2	18 500	1 200	60 000
3	17 200	1 500	50 000
4	15 500	1 700	30 000
5	14 000	1 800	10 000
6	12 200	2 200	5 000

计算每个阶段上的机器使用年数的可行值有一点难度. 图 9.6 给出了这个问题的网络表示. 在第 1 年年初, 我们有一台使用了 3 年的机器. 我们既可以更新它 (R), 也可以继续使用 1 年 (K). 在第 2 年年初, 如果那台机器已经在第一年得到更新, 那么这台新机器就已经使用 1 年了; 否则, 旧机器就使用 4 年了. 我们对第 2 年到第 4 年应用同样的逻辑, 假如运行了 1 年的机器在第 2, 3, 4 年年初进行更新的话, 在下一年年初这台更新了的机器就成了运行了 1 年的旧机器. 另外, 在第 4 年年初, 一台使用了 6 年的机器必须更新, 在第 4 年年末 (整个计划期末) 的时候, 我们将所有这些机器折旧 (S) 处理.

这个网络表明, 在第 2 年年初, 可能的机器运行年数为 1 年和 4 年. 到了第 3 年年初, 可能的机器运行年数为 1 年、2 年和 5 年. 到第 4 年年初, 可能的机器年龄分别为 1 年、2 年、3 年和 6 年.

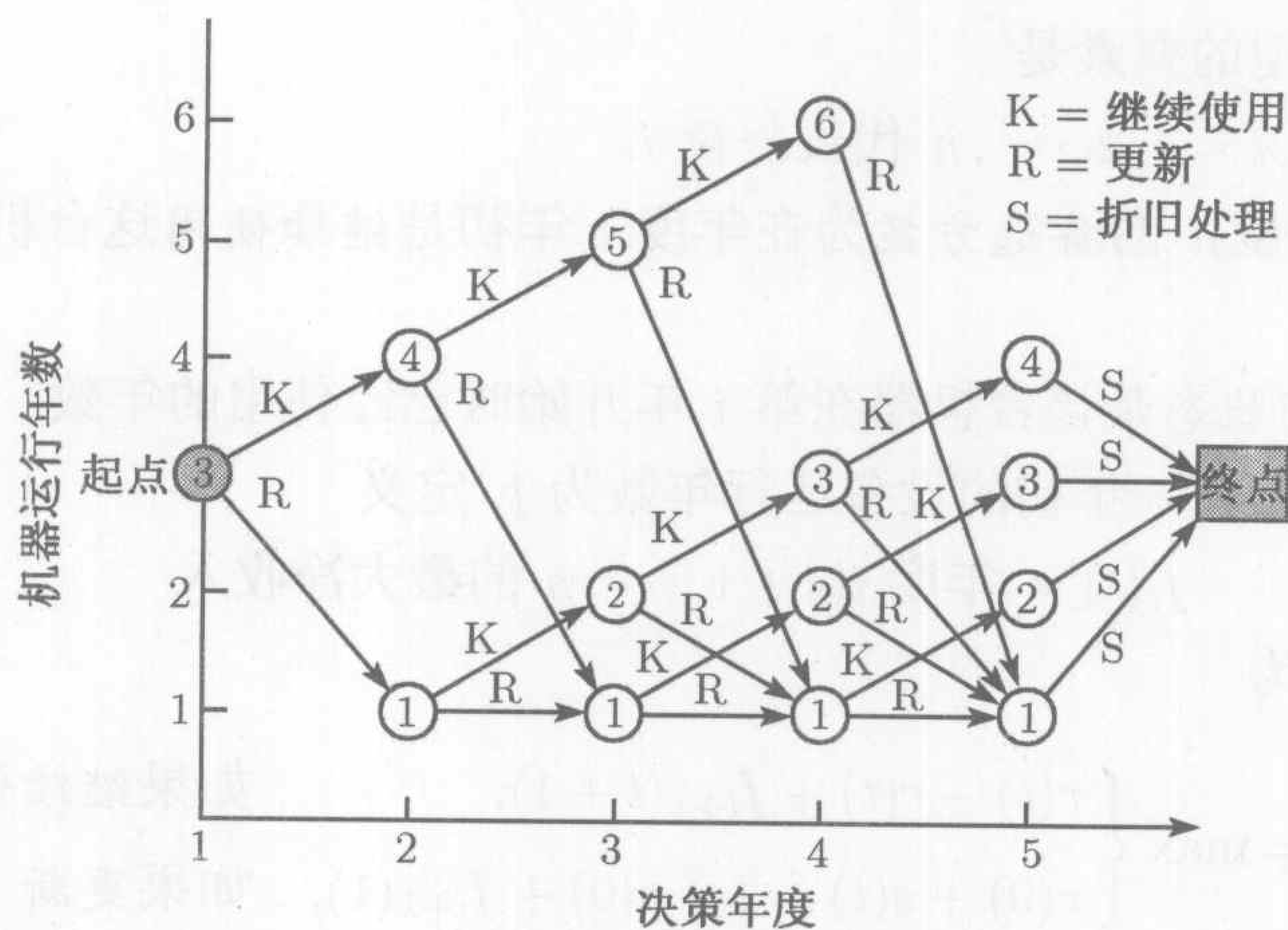


图 9.6　例 9.3-3 中作为决策年度函数的机器运行年数表示

求图 9.6 中网络的解, 等价于找出从第 1 年年初开始到第 4 年年末的一条最长的路径 (比如最大收益). 我们会利用表的形式来求解这个问题. 所有的货币值单位为 1 000 美元. 注意到, 如果在第 4 年 (即整个计划期末) 更新某台设备, 其收益将包括被更新机器的折旧现值 $s(t)$, 加上更新机器的折旧现值 $s(1)$.

阶段 4

t	K	R	最优解	
	$r(t)+s(t+1)-c(t)$	$r(0)+s(t)+s(1)-c(0)-I$	$f_4(t)$	决策结果
1	$19.0+60-0.6=78.4$	$20+80+80-0.2-100=79.8$	79.8	R
2	$18.5+50-1.2=67.3$	$20+60+80-0.2-100=59.8$	67.3	K
3	$17.2+30-1.5=45.7$	$20+50+80-0.2-100=49.8$	49.8	R
6	(必须更新)	$20+5+80-0.2-100=4.8$	4.8	R

阶段 3

t	K	R	最优解	
	$r(t)-c(t)+f_4(t+1)$	$r(0)+s(t)-c(0)-I+f_4(1)$	$f_3(t)$	决策结果
1	$19.0-0.6+67.3=85.7$	$20+80-0.2-100+79.8=79.6$	85.7	K
2	$18.5-1.2+49.8=67.1$	$20+60-0.2-100+79.8=59.6$	67.1	K
5	$14.0-1.8+4.8=17.0$	$20+10-0.2-100+79.8=19.6$	19.6	R

阶段 2

t	K	R	最优解	
	$r(t)-c(t)+f_3(t+1)$	$r(0)+s(t)-c(0)-I+f_3(1)$	$f_2(t)$	决策结果
1	$19.0-0.6+67.1=85.5$	$20+80-0.2-100+85.7=85.5$	85.5	K 或 R
4	$15.5-1.7+19.6=33.4$	$20+30-0.2-100+85.7=35.5$	35.5	R

阶段 1

t	K	R	最优解	
	$r(t)-c(t)+f_2(t+1)$	$r(0)+s(t)-c(0)-I+f_2(1)$	$f_1(t)$	决策结果
3	$17.2-1.5+35.5=51.2$	$20+50-0.2-100+85.5=55.3$	55.3	R

图 9.7 总结了最优解的求解结果. 在第 1 年年初, 由于 $t=3$, 最优决策是更新这台机器. 这样, 在第 2 年年初时, 这台新机器的年龄为 1, 并且在第 2 年年初 $t=1$ 要求这台机器要么继续使用, 要么更新. 如果更新了, 新机器在第 3 年年初时, 年龄为 1; 否则, 继续使用的机器年龄为 2. 继续这一过程, 直到第 4 年为止.

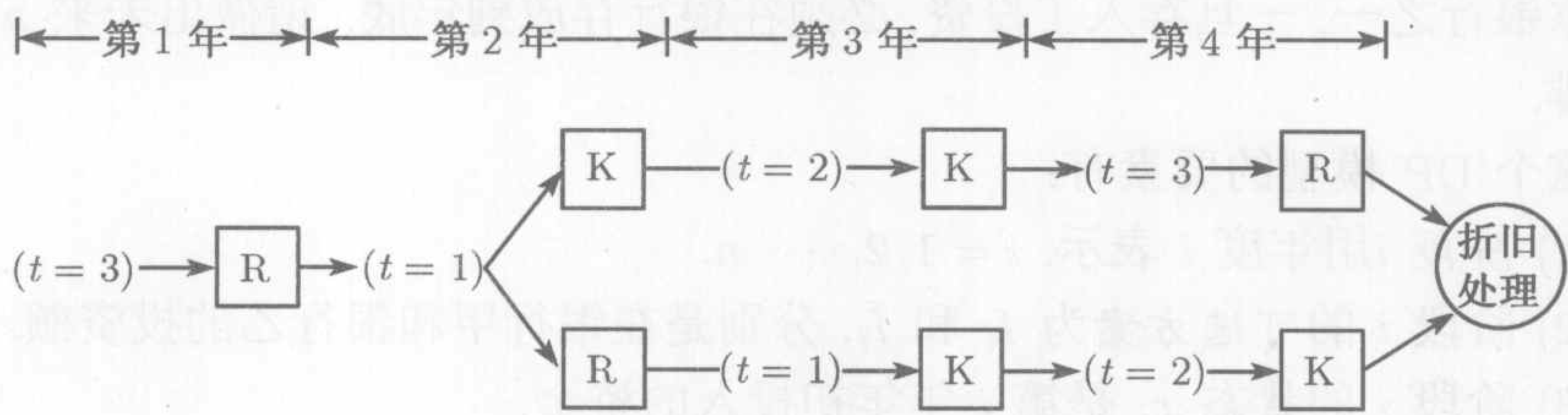


图 9.7 例 9.3-3 的解

从第 1 年开始的可能方案的最优策略是 (R,K,K,R) 和 (R,R,K,K), 总费用为 \$55 300.

习题 9.3C

1. 针对下列每种情况, 做出网络图, 求出例 9.3-3 模型的最优解:

(a) 第 1 年年初时, 机器已经使用了 2 年. (b) 第 1 年年初时, 机器刚使用了 1 年.

(c) 第 1 年年初时, 是刚买的新机器.

***2.** 我有个 13 岁的儿子, 他为 10 个顾客修剪草坪, 一年里为每个顾客剪草 3 次, 每次挣 \$50. 他刚花 \$200 买了台新的剪草机, 剪草机第 1 年的维修和运行费用为 \$120, 以后每年增加 20%. 运行了 1 年的二手剪草机价值 \$150, 随后每年降低 10%. 我儿子打算在他 16 岁前一直要做这项工作, 因此他想每 2 年就买台新的割草机, 因为他认为这样会更加划算, 他的依据是, 新剪草机的价格每年只涨 10%. 他这样决定正确吗?

3. Circle 农场想在未来 5 年内对一台已经使用了 2 年的拖拉机制定一个更新计划. 一台拖拉机必须至少使用 3 年, 但 5 年后必须淘汰. 拖拉机的现行价格为 \$40 000, 以后每年上涨 10%. 拖拉机的当前年度运行费是 \$1 300, 但预计每年增加 10%.

(a) 将该问题刻画成一个最短路径问题. (b) 给出相应的递归公式.

(c) 求出未来 5 年内的拖拉机最优更新策略.

4. 考虑 n 年期间的设备更新问题. 某种新设备价值 c 美元, 使用 t 年后的二手设备价值为: 如果 $n>t$, 则 $s(t)=n-t$, 否则为零. 年收益是设备年龄的函数, 表示为: 如果 $n>t$, 则 $r(t)=n^2-t^2$, 否则为零.

(a) 建立该问题的 DP 模型.

(b) 已知 $c=\$10\,000, n=5$, 且这是一台运行了 2 年的设备, 求出最优更新策略.

5. 假定设备使用了 1 年, 且 $n=4, c=\$6\,000, r(t)=\frac{n}{1+t}$, 求解第 4 题.

9.3.4 投资模型

假设你想在未来的 n 年内的每年年初进行投资, 各年投资额为 $P_1, P_2, \cdots, P_n$ 元. 你在两家银行有两个投资机会: 银行甲的利率为 r_1, 银行乙的利率是 r_2, 均为年利率. 为了鼓励存款, 这两家银行都对新的投资按投资额支付红利. 银行甲和银行乙第 i 年的红利百分比分别为 q_{i1} 和 q_{i2}. 到投资年度的年底银行支付红利, 你可以决定在任何一家银行进行下一年的投资. 这就意味着, 只有红利和新的资金投入到两家银行之一. 一旦存入了投资, 必须在银行存放到年底. 请做出未来 n 年的投资安排.

这个 DP 模型的要素有:

(1) 阶段 i用年度 i 表示, $i=1,2,\cdots,n$.

(2) 阶段 i 的*可选方案*为 I_i 和 $\bar{I}_i$, 分别是在银行甲和银行乙的投资额.

(3) 阶段 i 的状态 x_i 是第 i 年年初投入的资金.

注意到, 按定义 $\bar{I}_i=x_i-I_i$, 因此

$$\begin{aligned} x_1 &= P_1 \\ x_i &= P_i + q_{i-1,1}I_{i-1} + q_{i-1,2}(x_{i-1}-I_{i-1}) \\ &= P_i + (q_{i-1,1}-q_{i-1,2})I_{i-1} + q_{i-1,2}x_{i-1}, i=2,3,\cdots,n \end{aligned}$$

再投资额 x_i 只包含新投入的钱, 加上 $(i-1)$ 年度投资所得的红利.

定义 $f_i(x_i)=$ 给定 x_i 下的年度 $i, i+1, \cdots, n$ 的最优投资额, 接下来假定 I_i 和 (x_i-I_i) 为第 i 年分别在银行甲和银行乙投入的资金数, 定义 s_i 为第 n 年年底时的累计投资额. 令 $\alpha_k=(1+r_k), k=1,2$, 这个问题可以描述成

$$\max z = s_1 + s_2 + \cdots + s_n$$

其中

$$\begin{aligned} s_i &= I_i\alpha_1^{n+1-i} + (x_i-I_i)\alpha_2^{n+1-i} \\ &= (\alpha_1^{n+1-i}-\alpha_2^{n+1-i})I_i + \alpha_2^{n+1-i}x_i, \ i=1,2,\cdots,n-1 \\ s_n &= (\alpha_1+q_{n1}-\alpha_2-q_{n2})I_n + (\alpha_2+q_{n2})x_n \end{aligned}$$

s_n 中加入 q_{n1} 项和 q_{n2} 项, 是因为年度 n 的红利是投资金额最终累计的一部分.

得出的后向 DP 递归公式为

$$\begin{aligned} f_i(x_i) &= \max_{0\leqslant I_i\leqslant x_i}\{s_i + f_{i+1}(x_{i+1})\}, \quad i=1,2,\cdots,n-1 \\ f_{n+1}(x_{n+1}) &\equiv 0 \end{aligned}$$

如前面所述, x_{i+1} 是用 x_i 定义的.

例 9.3-4

假设你想要今年投入 \$4 000, 在第 2 年到第 4 年年初分别投资 \$2 000. 银行甲的年利率是 8%, 未来 4 年里的红利分别是 1.8%、1.7%、2.1% 和 2.5%. 银行乙的年利率比银行甲低 0.2%, 但红利高 0.5%. 我们的目标是使得第 4 年年底的累计资金额达到最大.

利用前面使用的符号, 我们有

$$\begin{aligned}
&P_1 = \$4\,000, P_2 = P_3 = P_4 = \$2\,000\\
&\alpha_1 = (1+0.08) = 1.08\\
&\alpha_2 = (1+0.078) = 1.078\\
&q_{11} = 0.018, q_{21} = 0.017, q_{31} = 0.021, q_{41} = 0.025\\
&q_{12} = 0.023, q_{22} = 0.022, q_{32} = 0.026, q_{42} = 0.030
\end{aligned}$$

阶段 4

$$f_4(x_4) = \max_{0 \leqslant I_4 \leqslant x_4} \{s_4\}$$

其中

$$s_4 = (\alpha_1 + q_{41} - \alpha_2 - q_{42})I_4 + (\alpha_2 + q_{42})x_4 = -0.003I_4 + 1.108x_4$$

函数 s_4 对于 I_4 在区间 $0 \leqslant I_4 \leqslant x_4$ 上是线性的, 由于 I_4 的系数是负的, 它在 $I_4 = 0$ 时达到最大值. 因此, 阶段 4 最优解情况为

状态	最优解	
	$f_4(x_4)$	I_4^*
x_4	$1.108x_4$	0

阶段 3

$$f_3(x_3) = \max_{0 \leqslant I_3 \leqslant x_3} \{s_3 + f_4(x_4)\}$$

其中

$$\begin{aligned}
s_3 &= (1.08^2 - 1.078^2)I_3 + 1.078^2 x_3 = 0.004\,32I_3 + 1.162\,1x_3\\
x_4 &= 2\,000 - 0.005I_3 + 0.026x_3
\end{aligned}$$

因此

$$\begin{aligned}
f_3(x_3) &= \max_{0 \leqslant I_3 \leqslant x_3} \{0.004\,32I_3 + 1.162\,1x_3 + 1.108(2\,000 - 0.005I_3 + 0.026x_3)\}\\
&= \max_{0 \leqslant I_3 \leqslant x_3} \{2\,216 + 0.001\,22I_3 + 1.190\,9x_3\}
\end{aligned}$$

状态	最优解	
	$f_3(x_3)$	I_3^*
x_3	$2\,216 + 1.190\,9x_3$	0

阶段 2

$$f_2(x_2) = \max_{0\leqslant I_2\leqslant x_2} \{s_2 + f_3(x_3)\}$$

其中

$$s_2 = (1.08^3 - 1.078^3)I_2 + 1.078^3 x_2 = 0.006\ 985 I_2 + 1.252\ 73 x_2$$
$$x_3 = 2\ 000 - 0.005 I_2 + 0.022 x_2$$

因此

$$\begin{aligned} f_2(x_2) =& \max_{0\leqslant I_2\leqslant x_2} \{0.006\ 985 I_2 + 1.252\ 73 x_2 + 2\ 216 \\ & +1.190\ 9(2\ 000 - 0.005 I_2 + 0.022 x_2)\} \\ =& \max_{0\leqslant I_2\leqslant x_2} \{4\ 597.8 + 0.001\ 030\ 5 I_2 + 1.278\ 93 x_2\} \end{aligned}$$

状态	最优解	
	$f_2(x_2)$	I_2^*
x_2	$4\ 597.8 + 1.279\ 96 x_2$	x_2

阶段 1

$$f_1(x_1) = \max_{0\leqslant I_1\leqslant x_1} \{s_1 + f_2(x_2)\}$$

其中

$$s_1 = (1.08^4 - 1.078^4)I_1 + 1.078^4 x_1 = 0.010\ 05 I_2 + 1.350\ 4 x_1$$
$$x_2 = 2\ 000 - 0.005 I_1 + 0.023 x_1$$

因此

$$\begin{aligned} f_1(x_1) =& \max_{0\leqslant I_1\leqslant x_1} \{0.010\ 05 I_1 + 1.350\ 4 x_1 + 4\ 597.8 \\ & +1.279\ 96(2\ 000 - 0.005 I_1 + 0.023 x_1)\} \\ =& \max_{0\leqslant I_1\leqslant x_1} \{7\ 157.7 + 0.003\ 65 I_1 + 1.379\ 84 x_1\} \end{aligned}$$

状态	最优解	
	$f_1(x_1)$	I_1^*
$x_1 = \$4\ 000$	$7\ 157.7 + 1.383\ 49 x_1$	\$4 000

向后递归并注意到, $I_1^* = 4\ 000, I_2^* = x_2, I_3^* = I_4^* = 0$, 我们得到

$$x_1 = 4\ 000$$
$$x_2 = 2\ 000 - 0.005 \times 4\ 000 + 0.023 \times 4\ 000 = \$2\ 072$$
$$x_3 = 2\ 000 - 0.005 \times 2\ 072 + 0.022 \times 2\ 072 = \$2\ 035.22$$
$$x_4 = 2\ 000 - 0.005 \times 0 + 0.026 \times 2\ 035.22 = \$2\ 052.92$$

因此最优解可总结为

年度	最优解	决策方案	累计金额
1	$I_1^* = x_1$	在甲银行投资 $x_1 = \$4\,000$	$s_1 = \$5\,441.80$
2	$I_2^* = x_2$	在甲银行投资 $x_2 = \$2\,072$	$s_2 = \$2\,610.13$
3	$I_3^* = 0$	在乙银行投资 $x_3 = \$2\,035.22$	$s_3 = \$2\,365.13$
4	$I_4^* = 0$	在乙银行投资 $x_4 = \$2\,052.92$	$s_4 = \$2\,274.64$
累计投资$= f_1(x_1) = 7\,157.7 + 1.383\,49(4\,000) = \$12\,691.66(= s_1 + s_2 + s_3 + s_4)$			

习题 9.3D

1. 设 $r_1 = 0.085, r_2 = 0.08$, 另外假定 $P_1 = \$5\,000$, $P_2 = \$4\,000$, $P_3 = \$3\,000$, $P_4 = \$2\,000$, 求解例 9.3-4.

2. 某投资者有初始本金 \$10 000, 每年年底, 他必须决定花掉多少, 有多少钱要投资在定期账户上. 投资每 1 美元钱到年底能得到 $\alpha = \$1.09$, 而在任何一年里每花掉 y 美元钱能得到的享受相当于拥有 $\sqrt{y}$ 美元. 对于 5 年时间跨度, 用动态规划方法求解该问题.

3. 某农场主拥有 k 只羊. 在每年年底, 必须决定有多少只卖掉, 多少只要继续养. 年度 i 卖掉一只羊的利润为 P_i, 第 i 年留下来的羊到了第 $(i+1)$ 年时数目就会翻一倍. 该农场主计划到 n 年底卖掉所有的羊.

*(a) 给出该问题的通用递归公式.

(b) 对于 $n = 3$ 年, $k = 2$ 只羊, $p_1 = \$100, p_2 = \$130, p_3 = \$120$, 求解这个问题.

9.3.5 库存模型

动态规划在库存控制方面有着重要的应用, 第 10 章和第 15 章介绍了一些这类的应用问题, 第 10 章的模型是确定性的, 而第 15 章的模型是不确定性的.

9.4 维度问题

在所有我们介绍过的 DP 模型中, 任一阶段下的状态都是用某个单一因素表达的. 例如在背包模型中 (9.3.1 节), 唯一的限制条件是物品的重量. 实际上, 也可以把背包的容积作为另一个限制条件, 在这种情况下, 我们称一个阶段下的状态是二维的, 因为它由两个因素组成: 重量和容积.

状态变量数的增加也增加了每个阶段的计算量, 特别是 DP 表的计算量明显增加了, 因为每个表的行数对应于状态变量的所有可能的组合. 文献中有时把这种计算的难度认为是由**维度**引起的.

我们选择了下面的例子, 来展现这一维度问题, 也为了说明线性规划与动态规划之间的关系.

例 9.4-1

Acme 制造公司生产两种产品, 制造过程的日生产时间为 430 分钟. 每生产 1 件产品 1 需要 2 分钟, 每件产品 2 需要 1 分钟. 产品 1 的产量没有限制, 但产品 2

的最大日需求量是 230 件. 产品 1 的单件利润为 \$2, 产品 2 为 \$5. 请用动态规划方法求出最优解.

这个问题可表示成以下线性规划问题:

$$\begin{aligned} \max\ z &= 2x_1 + 5x_2 \\ \text{s.t.}\quad 2x_1 + x_2 &\leqslant 430 \\ x_2 &\leqslant 230 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$

DP 模型的要素是

(1) 阶段 i对应于产品 $i, i = 1, 2$.

(2) 可选方案x_i 就是产品 i 的数量, $i = 1, 2$.

(3) 状态(v_2, w_2) 代表阶段 2 所用的资源 1 和资源 2 的量 (生产时间和需求上限).

(4) 状态(v_1, w_1) 代表阶段 1 和阶段 2 所用的资源 1 和资源 2 的量 (生产时间和需求上限).

阶段 2 定义 $f_2(v_2, w_2)$ 为阶段 2(产品 2) 在状态 (v_2, w_2) 下的最大利润, 则

$$f_2(v_2, w_2) = \max_{\substack{0 \leqslant x_2 \leqslant v_2 \\ 0 \leqslant x_2 \leqslant w_2}} \{5x_2\}$$

因此, 当 $x_2 = \min\{v_2, w_2\}$ 时取得 $\max\{5x_2\}$, 并且阶段 2 的解为

状态	最优解	
	$f_2(v_2, w_2)$	x_2
(v_2, w_2)	$5\min\{v_2, w_2\}$	$\min\{v_2, w_2\}$

阶段 1

$$\begin{aligned} f_1(v_1, w_1) &= \max_{0 \leqslant 2x_1 \leqslant v_1} \{2x_1 + f_2(v_1 - 2x_1, w_1)\} \\ &= \max_{0 \leqslant 2x_1 \leqslant v_1} \{2x_1 + 5\min(v_1 - 2x_1, w_1)\} \end{aligned}$$

对阶段 1 的优化需要解一个极小化极大问题. 对当前问题, 我们设 $v_1 = 430$, $w_1 = 230$, 得到 $0 \leqslant 2x_1 \leqslant 430$. 由于 $\min(430 - 2x_1, 230)$ 是两条交叉线的下包络 (请验证!), 因此得到

$$\min(430 - 2x_1, 230) = \begin{cases} 230, & 0 \leqslant x_1 \leqslant 100 \\ 430 - 2x_1, & 100 \leqslant x_1 \leqslant 215 \end{cases}$$

并且

$$\begin{aligned} f_1(430, 230) &= \max_{0 \leqslant x_1 \leqslant 215} \{2x_1 + 5\min(430 - 2x_1, 230)\} \\ &= \max_{x_1} \begin{cases} 2x_1 + 1\,150, & 0 \leqslant x_1 \leqslant 100 \\ -8x_1 + 2\,150, & 100 \leqslant x_1 \leqslant 215 \end{cases} \end{aligned}$$

可以用图形证实：当 $x_1 = 100$ 时, $f_1(430, 230)$ 达到最优值. 因此有

状态	最优解	
	$f_1(v_1, w_1)$	x_1
$(430, 230)$	1 350	100

为求出 x_2 的最优值, 注意到

$$v_2 = v_1 - 2x_1 = 430 - 200 = 230$$

$$w_2 = w_1 - 0 = 230$$

结果

$$x_2 = \min(v_2, w_2) = 230$$

整个最优解为

$$x_1 = 100 \text{ 件},\ x_2 = 230 \text{ 件},\ z = \$1\,350$$

习题 9.4A

1. 用动态规划求解下列各题.

(a) $\max\ z = 4x_1 + 14x_2$

s.t. $2x_1 + 7x_2 \leqslant 21$

$7x_1 + 2x_2 \leqslant 21$

$x_1, x_2 \geqslant 0$

(b) $\max\ z = 8x_1 + 7x_2$

s.t. $2x_1 + x_2 \leqslant 8$

$5x_1 + 2x_2 \leqslant 15$

$x_1, x_2 \geqslant 0$ 且为整数

(c) $\max\ z = 7x_1^2 + 6x_1 + 5x_2^2$

s.t. $x_1 + 2x_2 \leqslant 10$

$x_1 - 3x_2 \leqslant 9$

$x_1, x_2 \geqslant 0$

2. 在例 9.3-1 的 n 件物品背包问题中, 假设重量和容积上限分别为 W 和 V. 给定 w_i, v_i, r_i 为每单位物品 i 的重量、容积和收入, 写出该问题的 DP 后向递归公式.

参 考 文 献

Bertsekas, D., *Dynamic Programming: Deterministic and Stochastic Models*, Prentice Hall, Upper Saddle River, NJ, 1987.

Denardo, E., *Dynamic Programming Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, 1982.

Dreyfus, S., and A. Law, *The Art and Theory of Dynamic Programming*, Academic Press, New York, 1977.

Sntedovich, M., *Dynamic Programming*, Marcel Dekker, New York, 1991.

第 10 章　确定性库存模型

本章导读　库存模型用来确定企业为保证正常生产所必需持有的某种商品库存水平. 这种决策的基础是利用这样一种模型, 它能够在库存过剩造成的资金占用与库存不足造成的惩罚损失之间达到平衡. 影响解的主要因素是需求的性质: 是确定性的, 还是随机性的. 在实际中, 需求往往是随机性的, 但在某些情况下, 确定性的近似也是可以接受的. 本章介绍确定性的库存模型. 随机库存模型在第 15 章中介绍.

由于库存问题的复杂性, 不可能建立一种一般性的模型来解决所有的问题. 本章介绍不同情形下的一些代表性模型. 在学习不同的模型时, 应该注意到, 问题的求解使用了不同的算法, 包括微积分、线性、非线性和动态规划方法. 不管用什么方法求解模型, 你应该牢记的是, 任何库存模型无非得出两个基本结果: 订货多少和什么时候订货.

对某些模型的计算可能会很麻烦. 为了方便读者, 我们在本章给出了一些 Excel 电子表格、规划求解和 AMPL 模型, 可用来做实验 (如改变模型参数进行灵敏度分析), 或用来检查你的解题计算结果.

本章包含了 8 个例题、1 个规划求解模型、1 个 AMPL 模型、4 个 Excel 电子表格、33 个节末习题以及 3 个案例. 这些案例都在本书 (下册) 附录 E 中, AMPL/Excel/Solver/TORA 程序放在下载资料的 ch10Files 文件夹下.

10.1　一般库存模型

库存问题涉及定期发出和接收规定数量的订货. 从这一点来看, **库存策略** (inventory policy) 需要回答两个问题:

(1) 订多少货?　　(2) 什么时候订货?

回答这些问题的依据, 是使得以下库存费用函数达到最小.

$$\text{库存总费用} = \text{采购费} + \text{订货费} + \text{存货费用} + \text{缺货损失}$$

(1) **采购费**指库存货物的单价. 有时候, 如果订货量超过某个数目, 可能得到折扣, 这也成为决定订多少货的一个因素.

(2) **订货费**表示每进一次货所要支付的固定费用, 与每次的订货量无关. 根据已知需求增加每次的订货量会减少总的订货费, 但会增加平均库存水平, 造成库存

费用增加. 另一方面, 减少订货量会增加订货次数, 从而增加相应的订货费. 库存费用模型要在这两种费用之间进行平衡.

(3) *存货费用*表示维持库存的费用, 包括占用资金的利息、储存成本、维护和管理费用.

(4) *缺货损失*指发生缺货情况下所导致的惩罚费用, 包括可能的收入损失, 以及对客户丧失信誉的主观费用.

一个库存系统可能要根据**周期盘点**(periodic review)(如每周或每月订一次货), 每个周期开始时发出新订单; 也可能根据**连续盘点** (continuous review), 当库存量下降到某个水平 [称为**订货点** (reorder point)] 时, 再发出新订单. 周期盘点的例子如加油站, 每周开始新的供货; 而连续盘点出现零售店, 只有当它们的货架商品 (如化妆品) 数量低于某个水平时, 才进行补货.

10.2 需求在库存模型中的作用

一般来说, 库存模型的解析复杂性取决于对某种货物的需求是确定性的还是随机性的. 不论在哪一种情况下, 需求都有可能随着时间发生改变. 例如, 家庭取暖所用天然气的消耗量, 是一年内不同时间的函数, 隆冬季节达到最大, 春季和夏季月份里逐渐减少. 虽然这种季节性的变化每年都会重演, 但各个年份同一月份的消耗量还是有所变化, 例如根据天气的情况而不同.

在实际情况中, 库存模型中的需求模式可分为 4 类:

(1) 确定性的, 与时间无关的常数 (静态).

(2) 确定性的, 但随时间变化 (动态).

(3) 随机性的, 时间平稳的.

(4) 随机性的, 时间非平稳的.

这种分类假定我们关于未来需求的数据是可用的.

从建立库存模型的角度, 第 1 类最简单, 第 4 类最复杂. 然而在实际情况下, 第 1 类最少发生, 而第 4 类最普遍. 实际上, 我们需要在模型简化和模型精确性方面做出某种平衡, 即我们既不想用最简单但不反映现实的模型, 也不想用虽能反映现实但却分析不出来的复杂模型.

那么怎么决定对需求的某种近似是可接受的呢? 可以从计算某个周期比如每个月的消费量的平均值和标准差入手, 用变异系数 $V = \frac{\text{标准差}}{\text{平均值}} \times 100$ 并根据下面的原则来确定需求的性质:[①]

① 变异系数 V 度量数据在平均值附近的相对变化或分散情况. 一般地, V 的值越大, 表示使用平均值作为月需求量近似值时的不确定性越高. 对于确定性的需求, $V = 0$, 因为相应的标准差为零.

(1) 假如平均月需求量对于所有月份都“近似于”某个常数, 而且 V 比较小 ($< 20\%$), 则可以认为这个需求是确定性的, 并且是一个常数, 它的值等于所有月份需求量的平均值.

(2) 如果平均月需求量对于不同月份变化较大, 但 V 比较小, 则可以认为这个需求是确定性的, 但却是随时间变化的.

(3) 如果在第一种情况下, V 很大 ($> 20\%$) 但是接近于常数, 则这种需求是随机性的和平稳的.

(4) 剩下的最后一种情况是随机性的非平稳需求, 其平均值和变异系数随时间的变化都非常大.

在第 3 种和第 4 种情况下, 常常需要其他数据来确定相应的概率分布.

例 10.2-1

表 10.1 给出了 10 年内 (1990~1999) 农村居民 (1 月到 12 月) 液化天然气的月消费量数据. 无论什么时候只要居民提出要求, 天然气供应公司就会开卡车上门, 为住户灌气, 由住户决定灌气的时间和多少.

从库存建模的角度看, 我们完全可以假定, 每个月代表一个决策周期, 其间住户发出一个订货单. 然而我们主要关心的是, 如何分析需求的特性.

表 10.1

年	天然气消费量 (立方英尺)											
	1 月	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月
1990	100	110	90	70	65	50	40	42	56	68	88	95
1991	110	125	98	80	60	53	44	45	63	77	92	99
1992	90	100	88	79	56	57	38	39	60	70	82	90
1993	121	130	95	90	70	58	41	44	70	80	95	100
1994	109	119	99	75	68	55	43	41	65	79	88	94
1995	130	122	100	85	73	58	42	43	64	75	80	101
1996	115	100	103	90	76	55	45	40	67	78	98	97
1997	130	115	100	95	80	60	49	48	64	85	96	105
1998	125	100	94	86	79	59	46	39	69	90	100	110
1999	87	80	78	75	69	48	39	41	50	70	88	93
平均值	111.7	110	95	82.5	69.6	55.3	42.7	42.2	62.8	77.2	90.7	98
标准差	15.54	15.2	7.5	7.99	7.82	3.95	3.4	2.86	6.09	6.91	6.67	6
$V(\%)$	13.91	13.8	7.9	9.68	11.24	7.13	7.96	6.78	9.69	8.95	7.35	6.1

观察表 10.1 中的平均值和变异系数 V, 我们可得出两个结果:

(1) 平均消费量是动态 (非常数) 的, 因为它显示出, 冬季月份的平均消费量比夏季月份相对要高些.

(2) 变异系数 V 较小 ($< 15\%$), 因此可以认为月需求量是近似于确定性的.

因此这两个结果决定了我们可以建立一种月消费量 (近似) 确定性但随时间变化的

库存模型.

10.3　静态经济订货量 (EOQ) 模型

本节给出对于静态 (常数) 需求的 3 种不同的经济订货量模型. 这些模型的特点是简单并且易于分析.

10.3.1　经典 EOQ 模型

最简单的库存模型中, 需求增长率为常数, 即时补货, 而且没有缺货. 定义

$y =$ 订货量 (件数)

$D =$ 需求率 (单位时间货物件数)

$t_0 =$ 订货周期长度 (时间单位数)

库存水平服从图 10.1 显示的模式. 当库存下降到 0 水平时, 即时发出并得到一个 y 单位货物的订货. 库存以常数需求率 D 减少直至完全耗尽. 这种模式的订货周期为

$$t_0 = \frac{y}{D} \text{ 个时间单位}$$

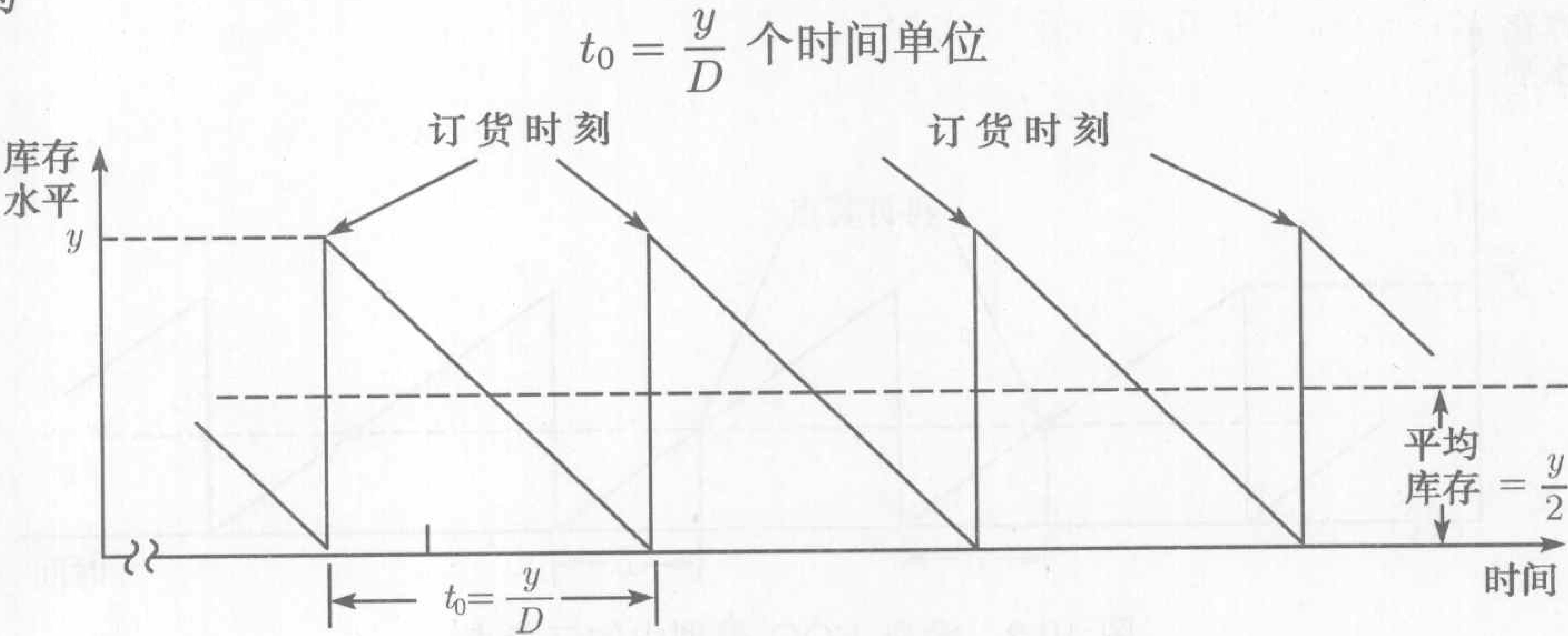


图 10.1　经典 EOQ 模型中的库存模式

该费用模型带有两个费用参数.

$K =$ 发出一份订单相应的订货费 (美元/订单)

$h =$ 存货费 [美元/(单位库存货物 · 单位时间)]

已知平均库存水平为 $\frac{y}{2}$, 每单位时间总费用 (TCU) 可按如下公式计算:

$$\begin{aligned}
TCU(y) &= \text{每单位时间的订货费} + \text{每单位时间的存货费} \\
&= \frac{\text{订货费} + \text{每周期 } t_0 \text{ 的存货费}}{t_0} \\
&= \frac{K + h\left(\frac{y}{2}\right)t_0}{t_0} \\
&= \left(\frac{K}{\frac{y}{D}}\right) + h\left(\frac{y}{2}\right)
\end{aligned}$$

订货量 y 的最优值由 $TCU(y)$ 关于 y 的最小值得出. 假设 y 是连续的, 找到 y 的最优值的必要条件为

$$\frac{\mathrm{d}TCU(y)}{\mathrm{d}y} = -\frac{KD}{y^2} + \frac{h}{2} = 0$$

由于 $TCU(y)$ 是凸的, 这个条件也是充分的.

该方程的解给出了 EOQ y^*:

$$y^* = \sqrt{\frac{2KD}{h}}$$

因此, 上述模型的最优库存策略就是

$$\text{每 } t_0^* = \frac{y^*}{D} \text{ 时间单位订 } y^* = \sqrt{\frac{2KD}{h}} \text{ 单位货物}$$

实际上, 一个新的订单并不一定是在订货时立即接收到货, 而是在发出订单与收到货物之间存在一个正的**提前时间** (lead time) L, 如图 10.2 所示. 在这种情况下, 当库存水平降到 LD 单位时, 就会有一个**再订货点** (reorder point).

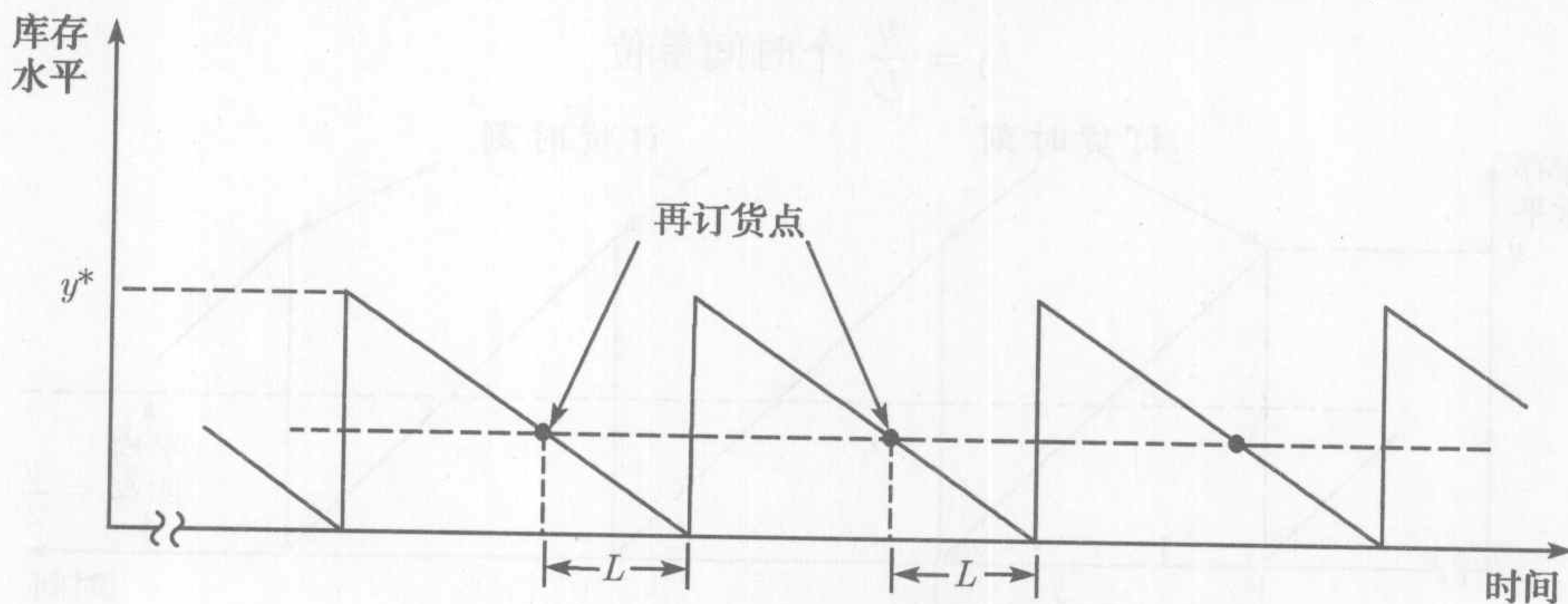


图 10.2 经典 EOQ 模型中的订货点

图 10.2 中假设, 提前时间 L 小于订货周期长度 t_0^*, 但通常并不都是这样. 为了处理这一情况, 定义**有效提前时间** (effective lead time):

$$L_e = L - nt_0^*$$

其中 n 为不超过 $\frac{L}{t_0^*}$ 的最大整数. 这个结果是成立的, 因为在每个 t_0^* 的 n 个周期以后, 这种库存情况就好像是, 从发出订单到收到下一批货物之间的间隔为 L_e, 因此再订货点发生在 L_eD 单位, 并且库存策略可以重新叙述为

每当库存水平下降到 L_eD 单位时, 则发出订货量 y^* 的订单

例 10.3-1

某大学校园要更换霓虹灯, 每天更换 100 个灯管, 后勤工厂定期发出订货. 假定发出一份采购订单的订货费为 \$100, 在仓库里存放一个霓虹灯管的费用大约为

每天 \$0.02. 订货与接到货之间的提前时间为 12 天, 请给出这批霓虹灯管订货的最优库存策略.

根据问题的数据, 我们有

$$D = \text{每天 } 100 \text{ 个}$$
$$K = \text{每份订单 } \$100$$
$$h = \text{每天每个灯管 } \$0.02$$
$$L = 12 \text{ 天}$$

因此有

$$y^* = \sqrt{\frac{2KD}{h}} = \sqrt{\frac{2 \times \$100 \times 100}{0.02}} = 1\,000 \text{个灯管}$$

相应的周期长度为

$$t_0^* = \frac{y^*}{D} = \frac{1\,000}{100} = 10\text{天}$$

由于提前时间 $L = 12$ 天超过了周期长度 $t_0(=10$ 天), 必须计算 L_e. L 中的整数周期数为

$$n = \left(\text{最大整数} \leqslant \frac{L}{t_0^*}\right) = \left(\text{最大整数} \leqslant \frac{12}{10}\right) = 1$$

因此

$$L_e = L - nt_0^* = 12 - 1 \times 10 = 2 \text{ 天}$$

当库存水平下降到

$$L_eD = 2 \times 100 = 200 \text{ 个霓虹灯管}$$

时就需要再订货. 霓虹灯管订货的最优库存策略为

一旦库存水平下降到 200 个灯管时就订 1 000 个.

这一库存策略相应的日库存费用为

$$TCU(y) = \frac{K}{\left(\frac{y}{D}\right)} + h\left(\frac{y}{2}\right) = \frac{\$100}{\left(\frac{1\,000}{100}\right)} + \$0.02\left(\frac{1\,000}{2}\right) = \$20\ /\text{天}$$

Excel 程序

我们设计了一个计算 EOQ 的程序 exelEOQ.xls. 这个模型用来求解习题 10.3A 第 10 题中给出的带有缺货和同时进行生产–消费活动的一般性 EOQ 问题, 上面介

绍的模型是这类问题的一个特殊情况. 这个模型还可以解 10.3.2 节给出的分段价格情形. 在对例 10.3-1 使用本程序时, 请在单元 C3:C5、C8 和 C10 中输入 −1, 表示对应的数据不存在, 如图 10.3 所示.

	B	C	D
1	General Economic Order Quantity (EOQ)		
2	Input data: Enter -1 in column C if data element does not apply		
3	Item cost, c1 =	-1	
4	Qty discount limit, q =	-1	
5	Item cost, c2 =	-1	
6	Setup cost, K =	100	
7	Demand rate, D =	100	
8	Production rate, a =	-1	
9	Unit holding cost, h =	0.02	
10	Unit penalty cost, p =	-1	
11	Lead time, L =	12	
12	Model output results:		
13	Order qty, y* =	1000.00	
14	Shortage qty, w* =	0.00	
15	Reorder point, R =	200.00	
16	TCU(y*) =	20.00	
17	Purchase/prod. Cost =	0.00	
18	Setup cost/unit time =	10.00	
19	Holding cost /unit time =	10.00	
20	shortage cost/unit time =	0.00	
21	Optimal inventory policy: Order 1000.00 units when level drops to 200.00 units		
22	Model intermediate calculations:		
23	ym =	1000.00	
24	TCU1(ym)=	Not applicable	
25	Q-equation:	Not applicable	
26	Q =	Not applicable	
27	cycle length, t0 =	10.00	
28	Optimization zone =	Not applicable	
29	Effectice lead time, Le =	2.00	

图 10.3 例 10.3-1 的 Excel 解 (文件 excelEOQ.xls)

习题 10.3A

1. 针对下列每种情形, 不允许缺货, 从订货到接到货的提前时间为 30 天, 给出最优库存策略以及相应的日库存费用.

(a) $K = \$100, h = \$0.05, D = 30$ 个/天

(b) $K = \$50, h = \$0.05, D = 30$ 个/天

(c) $K = \$100, h = \$0.01, D = 40$ 个/天

(d) $K = \$100, h = \$0.04, D = 20$ 个/天

2. McBurger 汉堡包店每周一开始都要订瘦肉, 每星期需求量是 300 磅. 每份订单的固定费用为 \$20, 每磅肉每天的冷冻和储存费用是 \$0.03.

(a) 算出现行订货方式的每周库存费用.

(b) 给出 McBurger 应该采用的最优库存策略, 假定从订货到接收到货的提前时间为零.

3. 某公司库存有某种货物, 每天消费 50 件. 公司为每次订货支付订货费 \$20, 库存每件货物每周的费用为 \$0.35.

(a) 给出最优库存策略, 假定提前时间为一周.

(b) 求出每年最优订货次数 (一年按 365 天算).

4. 某公司的采购部门提出了下面两个库存策略:

策略 1 订货量 150 件, 再订货点为 50 件, 从订货到接收到货的提前时间为 10 天.

策略 2 订货量 200 件, 再订货点为 75 件, 从订货到接收到货的提前时间为 15 天. 每次订货的订货费为 \$20, 每件库存的存货费用为每天 \$0.02.

(a) 该公司应采用哪种优库存策略?

(b) 如果是你负责制定该公司的库存策略, 假定供应商要求 22 天的提前时间, 你的建议会是什么?

5. Walmark 商场把空的纸盒压缩, 打成箱准备回收. 商场每天能产生 5 箱, 在商场后院存放每箱的费用是每天 \$0.10. 有一家公司负责把这些纸盒箱运送到回收站, 收取固定费用 \$100 用来租装卸车, 外加运输费每箱 \$3. 请用图表示纸盒箱数随时间的变化, 并给出运送这些箱子到回收站的最优策略.

6. 某旅馆把毛巾送到外面的清洗店去洗. 旅馆每天有 600 条脏毛巾要洗, 清洗店定期上门来取这些脏毛巾, 并换成洗好的干净毛巾. 洗衣店每次取送服务收取上门费 \$81, 外加每条毛巾清洗费 \$0.60. 旅店每天存放一条脏毛巾的费用为 \$0.02, 存放每条干净毛巾每天的费用是 \$0.01. 旅店该如何使用清洗店的送货上门服务呢? (**提示**: 本题有两种库存物品. 随着脏毛巾的增多, 干净毛巾以同样的速率减少.)

7. (Lewis, 1996) 一家跨国公司的某雇员在美国贷了款, 要去在欧洲的子公司工作. 在这一年里, 该雇员在美国总共要还款 \$12 000(包括房贷和保险费), 均匀分摊在当年的各个月份. 该雇员为了还款, 可以在出发去欧洲之前把这笔总计还款存在美国银行里. 但是, 现在美国银行的利率相当低 (年息约 1.5%), 而欧洲银行的利率很高 (年息 6.5%). 海外汇款的费用为每笔 \$50. 请你确定从欧洲汇款到美国的最优策略, 并讨论如何实行这一方案. 阐述所有的假设条件.

8. 考虑某库存情形, 其中库存以速率 a 均匀地 (而不是即时地) 补充, 消耗速率为常数 D. 由于在补充库存期间也进行消耗, 必须使 $a > D$. 每次订货的订货费为 K, 每单位时间每件货物的储存费为 h. 如果 y 为订货量, 不允许缺货, 请说明

(a) 最大库存水平是 $y\left(1-\frac{D}{a}\right)$.

(b) 给定 y 的每单位时间的总费用为

$$TCU(y) = \frac{KD}{y} + \frac{h}{2}\left(1-\frac{D}{a}\right)y$$

(c) 经济订货量为

$$y^* = \sqrt{\frac{2KD}{h\left(1-\frac{D}{a}\right)}}, D < a$$

(d) 说明从 (c) 中的公式能导出即时补货情况下的 EOQ.

9. 某公司可自己生产某种产品, 也可向某个合同商购买. 假如自产, 每次启动机器的成本是 \$20, 生产速度是每天 100 个. 如果是向合同商购买, 每次订货费为 \$15. 不论自产还是购买, 每件产品的库存费用是每天 \$0.02. 公司对该产品的年使用量为 26 000 个. 假定不允许出现缺货, 公司是该购买还是该自产呢?

10. 在第 8 题中, 假设允许缺货, 每件缺货每单位时间的惩罚费用为 p 元.

(a) 假如 w 为库存周期内的最大缺货量, 证明

$$TCU(y,w)=\frac{KD}{y}+\frac{h\left\{y\left(1-\frac{D}{a}\right)-w\right\}^2+pw^2}{2\left(1-\frac{D}{a}\right)y}$$

$$y^*=\sqrt{\frac{2KD(p+h)}{ph\left(1-\frac{D}{a}\right)}}$$

$$w^*=\sqrt{\frac{2KDh\left(1-\frac{D}{a}\right)}{p(p+h)}}$$

(b) 说明 10.3.1 节中的 EOQ 结果可以从 (a) 的一般性公式中推出.

10.3.2　分段价格的 EOQ 模型

这种模型和 10.3.1 节中的模型唯一不同的是, 假如订货数量 y 超过某个给定的上限 q 时, 库存货物能够按折扣价购买. 从数学上, 单位采购价格 c 可表示为

$$c=\begin{cases}c_1, & 若\ y\leqslant q\\ c_2, & 若\ y>q\end{cases},\quad c_1>c_2$$

因此,

$$每单位时间的采购费用=\begin{cases}\dfrac{c_1y}{t_0}=\dfrac{c_1y}{\left(\frac{y}{D}\right)}=Dc_1, & y\leqslant q\\ \dfrac{c_2y}{t_0}=\dfrac{c_2y}{\left(\frac{y}{D}\right)}=Dc_2, & y>q\end{cases}$$

采用 10.3.1 节的符号, 每单位时间的总费用为

$$TCU(y)=\begin{cases}TCU_1(y)=Dc_1+\dfrac{KD}{y}+\dfrac{h}{2}y, & y\leqslant q\\ TCU_2(y)=Dc_2+\dfrac{KD}{y}+\dfrac{h}{2}y, & y>q\end{cases}$$

图 10.4 给出了函数 TCU_1 和 TCU_2 的图像. 因为这两个函数只相差一个常数量, 它们的最小值刚好都取在

$$y_m=\sqrt{\frac{2KD}{h}}$$

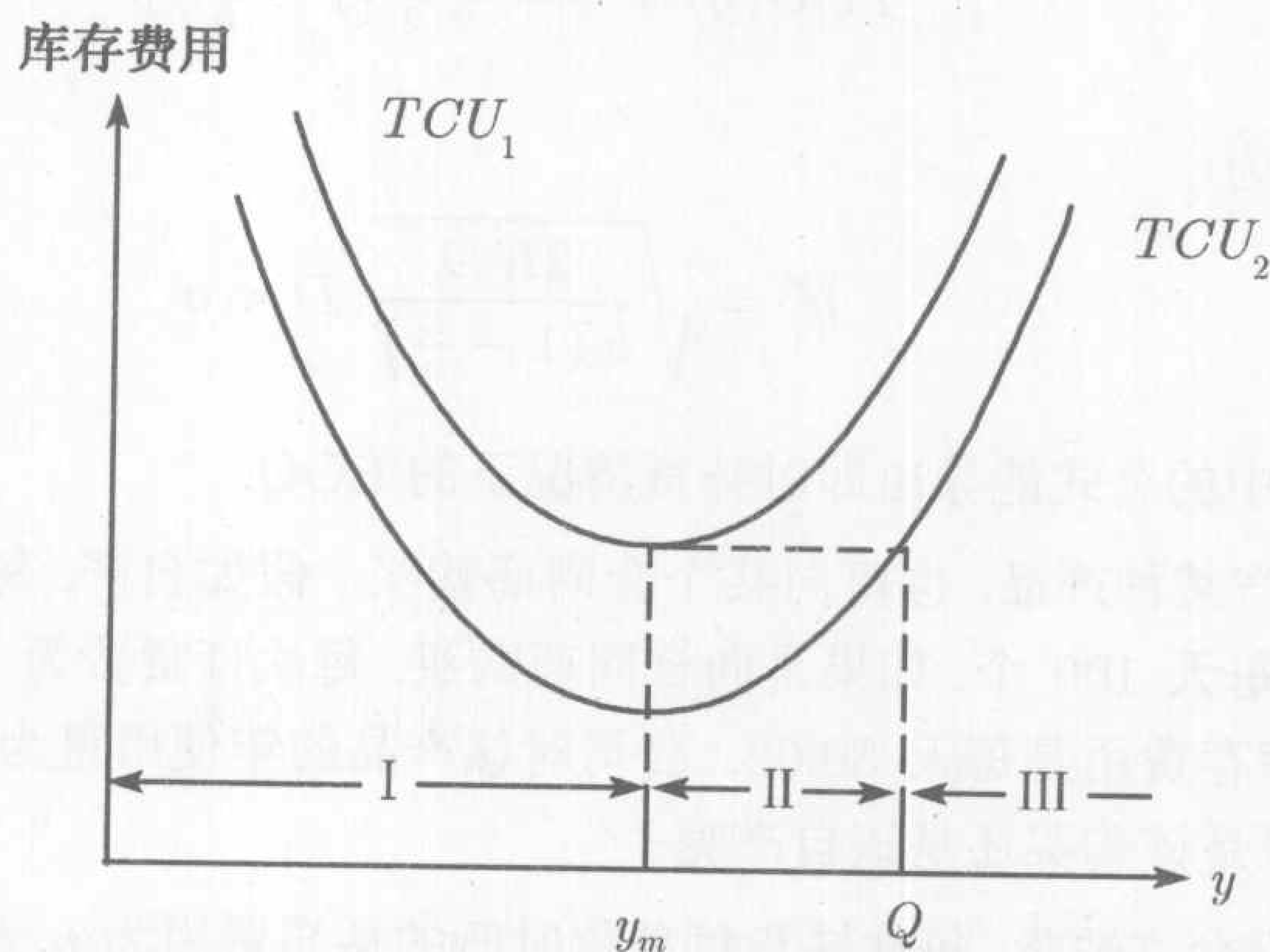


图 10.4　具有分段价格的库存费用函数

费用函数 $TCU(y)$ 从左边开始取 $TCU_1(y)$, 并在价格断点 q 处开始取 $TCU_2(y)$. 最优订货量 y^* 取决于价格断点 q 落在区域 I、II 还是 III, 图 10.4 中分别用 $(0, y_m)$, (y_m, Q), (Q, ∞) 给予了表示. $Q(> y_m)$ 的值可从以下公式得出:

$$TCU_2(Q) = TCU_1(y_m)$$

或

$$c_2 D + \frac{KD}{Q} + \frac{hQ}{2} = TCU_1(y_m)$$

简化为

$$Q^2 = \left(\frac{2(c_2 D - TCU_1(y_m))}{h}\right) Q + \frac{2KD}{h} = 0$$

图 10.5 显示出所需要的最优订货量 y^* 为

$$y^* = \begin{cases} y_m, & \text{若 } q \text{ 位于区间 I 或 III} \\ q, & \text{若 } q \text{ 位于区间 II} \end{cases}$$

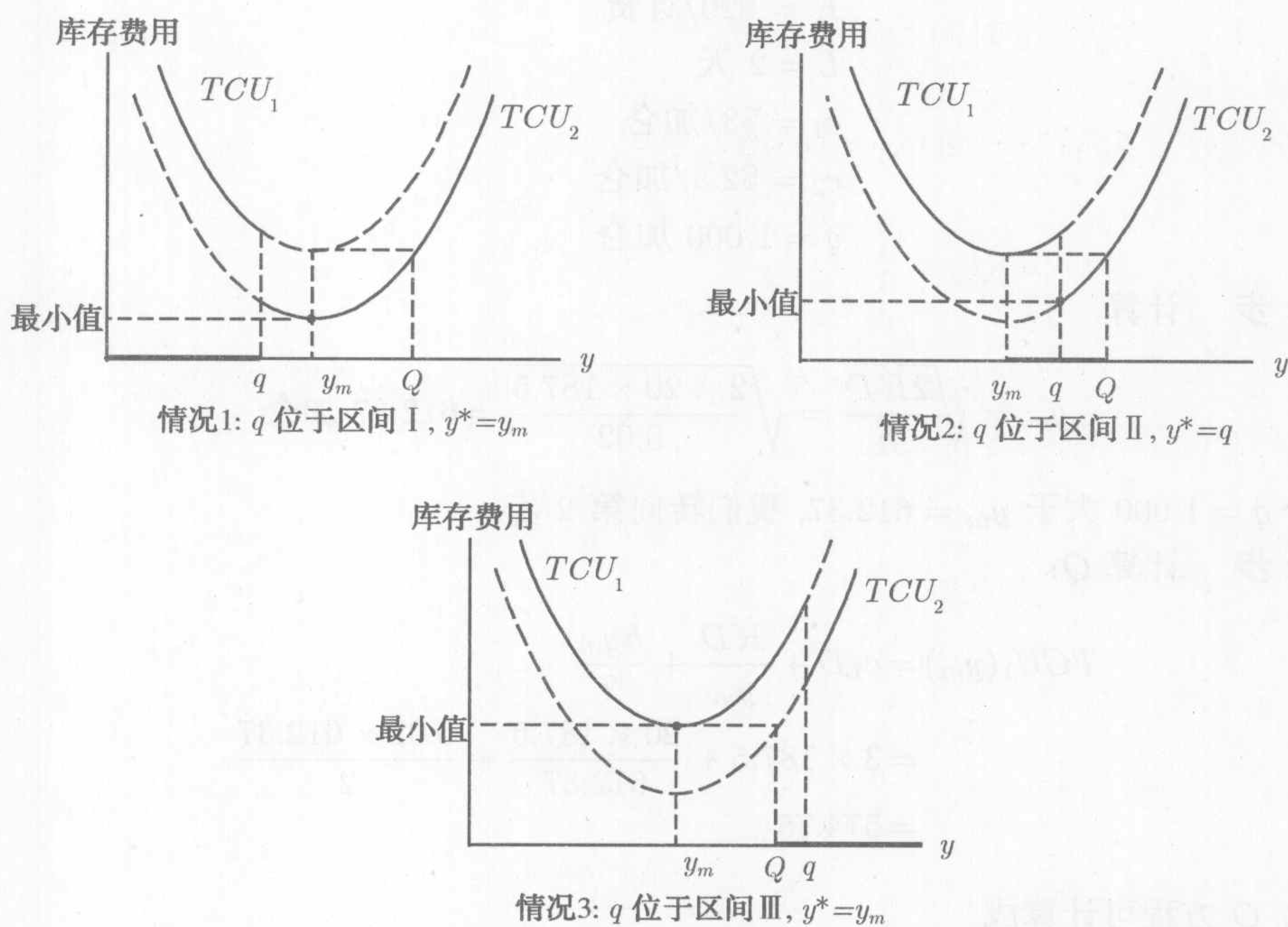


图 10.5 分段价格库存问题的最优解

计算 y^* 的步骤是

第 1 步 求出 $y_m = \sqrt{\frac{2KD}{h}}$. 如 q 位于区间 I, 则 $y^* = y_m$; 否则, 转到第 2 步.

第 2 步 根据 Q 方程求出 $Q(> y_m)$

$$Q^2+\left(\frac{2(c_2D-TCU_1(y_m))}{h}\right)Q+\frac{2KD}{h}=0$$

定义区间 II 和 III. 若 q 在区间 II 里, $y^*=q$; 否则, q 位于区间 III, 且 $y^*=y_m$.

例 10.3-2

LubeCar 公司专门给汽车快速换机油. 汽修间按桶购买机油, 每加仑 \$3. 如果 LubeCar 采购 1 000 加仑以上的话, 则每加仑的优惠价为 \$2.5. 汽修间每天可为 150 辆车提供服务, 每次换机油要用掉 1.25 加仑. LubeCar 储存成批油的费用是每加仑每天 \$0.02. 此外, 每次订货费为 \$20, 供货的提前时间有 2 天. 求出最优的库存策略.

每天机油的消耗量为

$$D=150\ \text{辆车/天}\ \times 1.25\ \text{加仑/车}\ =187.5\ \text{加仑/天}$$

我们还知道

$$\begin{aligned}
h&=\$0.02/(\text{天}\cdot\text{加仑})\\
K&=\$20/\text{订货}\\
L&=2\ \text{天}\\
c_1&=\$3/\text{加仑}\\
c_2&=\$2.5/\text{加仑}\\
q&=1\ 000\ \text{加仑}
\end{aligned}$$

第 1 步 计算

$$y_m=\sqrt{\frac{2KD}{h}}=\sqrt{\frac{2\times 20\times 187.5}{0.02}}=612.37\ \text{加仑}$$

由于 $q=1\ 000$ 大于 $y_m=612.37$, 我们转向第 2 步.

第 2 步 计算 Q:

$$\begin{aligned}
TCU_1(y_m)&=c_1D+\frac{KD}{y_m}+\frac{hy_m}{2}\\
&=3\times 187.5+\frac{20\times 187.5}{612.37}+\frac{0.02\times 612.37}{2}\\
&=574.75
\end{aligned}$$

因此, Q 方程可计算成

$$Q^2+\left(\frac{2\times(2.5\times 187.5-574.75)}{0.02}\right)Q+\frac{2\times 20\times 187.5}{0.02}=0$$

或

$$Q^2-10\ 599.74Q+375\ 000=0$$

由此得到 $Q = 10\ 564.25(> y_m)$, 因此

$$\text{区间 II} = (612.37, 10\ 564.25)$$

$$\text{区间 III} = (10\ 564.25, \infty)$$

因为 $q(= 1\ 000)$ 落在区间 II, 最优订货量为 $y^* = q = 1\ 000$ 加仑.

已知提前时间为 2 天, 再订货点为 $2D = 2 \times 187.5 = 375$ 加仑. 因此, 最优库存策略就是

当库存水平下降到 375 加仑时, 订货 1 000 加仑.

Excel 程序

Excel 程序 excelEOQ.xls 可用来求解上述带有折扣价的情形. 只要在电子表格输入数据区域 (C3:C11) 输入模型数据, 就可以直接计算这个模型. 如果输入数据有冲突, 就会显示适当的错误消息. 模型的输出给出最优库存策略, 以及问题的所有中间计算.

习题 10.3B

1. 考虑习题 10.3A 第 6 题中的旅馆清洗服务问题. 清洗一条脏毛巾的标准收费是 \$0.60, 但如果旅馆给清洗店至少 2 500 条毛巾的话, 清洗店清洗每条毛巾只收取 \$0.50. 旅馆应该利用这项打折服务吗?

*2. 某类货物的日消耗量是 30 件, 每天每件库存的费用为 \$0.05, 订货费 \$100. 假设不允许缺货, 如果一次购买量不超过 500 件时, 采购单价为 \$10, 否则按 \$8 计. 订货提前时间为 21 天, 请求出最优库存策略.

3. 某产品销售单价为 \$25, 但采购量超过 50 件的, 享受 10% 的优惠. 某公司每天使用该产品 20 件, 订货费每次 \$50, 每天每件存货费 \$0.30, 订货提前时间为 12 天. 该公司该利用这一优惠价吗?

4. 在第 3 题中, 请确定出价格优惠百分比的某个范围, 使得当批量订货超过 150 件时, 不会对公司产生经济上的好处.

5. 在本节讨论的库存模型中, 假设每件货物单位时间的存货费用当数量少于 q 时为 h_1, 否则为 h_2, $h_1 > h_2$. 说明如何确定经济订货批量.

10.3.3 带有储存上限的多货品 EOQ 模型

这类模型针对 $n(> 1)$ 种货品, 其中每种货品的库存波动符合图 10.1 的模式 (不允许缺货). 不同的是, 这些货品相互竞争有限的储存空间.

对于货品 $i, i = 1, 2, \cdots, n$, 定义

$$D_i = \text{需求率}$$

$$K_i = \text{订货费}$$

$$h_i = \text{单位时间每件存货费}$$

$$y_i = \text{订货量}$$
$$a_i = \text{每单件货品储存面积需求}$$
$$A = \text{所有 } n \text{ 种货品的最大可用储存面积}$$

在没有缺货的假设下, 该库存问题的数学模型可表示为

$$\min \quad TCU(y_1, y_2, \cdots, y_n) = \sum_{i=1}^{n}\left(\frac{K_i D_i}{y_i} + \frac{h_i y_i}{2}\right)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} a_i y_i \leqslant A$$

$$y_i > 0, \quad i = 1, 2, \cdots, n$$

要求解这个问题, 首先求出一个无约束问题的解:

$$y_i^* = \sqrt{\frac{2K_i D_i}{h_i}}, \quad i = 1, 2, \cdots, n$$

如果该解满足约束条件, 则我们得到了最优解. 否则, 必须使其满足约束.

在本书的前几个版本中, 我们采用了拉格朗日算法 (Lagrangian algorithm) 和试错计算方法 (trial-and-error calculation) 来找出符合约束的最优解. 现在我们有了功能强大的软件包 (如 AMPL 和 Excel 规划求解), 可以直接用非线性规划来求解, 如下面的例题所示.

例 10.3-3

以下数据描述 3 种库存货品.

货品 i	K_i($)	D_i(每天库存数)	h_i($)	a_i(ft^2)
1	10	2	0.30	1
2	5	4	0.10	1
3	15	4	0.20	1
可用的总储存面积 = 25 ft^2				

无约束最优值 $y_i^* = \sqrt{\frac{2K_iD_i}{h_i}}, i = 1, 2, 3$ 分别为 11.55, 20.00, 24.49 件. 这些值不满足库存约束

$$y_1 + y_2 + y_3 \leqslant 25$$

因此, 该问题可以按照下面的介绍, 用 Excel 规划求解或 AMPL 作为非线性规划来求解. Excel 规划求解模型必须要做些调整, 以适合问题的规模. AMPL 模型可以用于任意多的货品数量, 只需要改一下输入数据.

最优解为 $y_1^* = 6.34$ 件, $y_2^* = 7.09$ 件, $y_3^* = 11.57$ 件, 费用 = \$13.62/天

Excel 规划求解程序

图 10.6 显示了如何用 Excel 规划求解作为非线性规划程序求解例 10.3-3 (文件名 solverConstrEOQ.xls), 图中给出了程序所用的公式和规划求解参数. 如大多

数非线性规划程序一样, 开始必须给出初始解值 (本程序第 9 行, $y_1 = y_2 = y_3 = 1$). 必须给出非零初始值是因为, 目标函数包括被 y_i 所除的项. 其实, 用 $K_iD_i/(y_i+\Delta)$ 代替 K_iD_i/y_i 可能更好一些, 其中 $\Delta > 0$, 是一个非常小量, 用来在迭代中防止被零除. 一般来说, 找出一个 (局部最优) 解之前需要不同的初始值. 图下方的这个最优解是全局的, 是因为目标函数和约束方程具有非常好的表现 (凸目标函数和凸解空间).

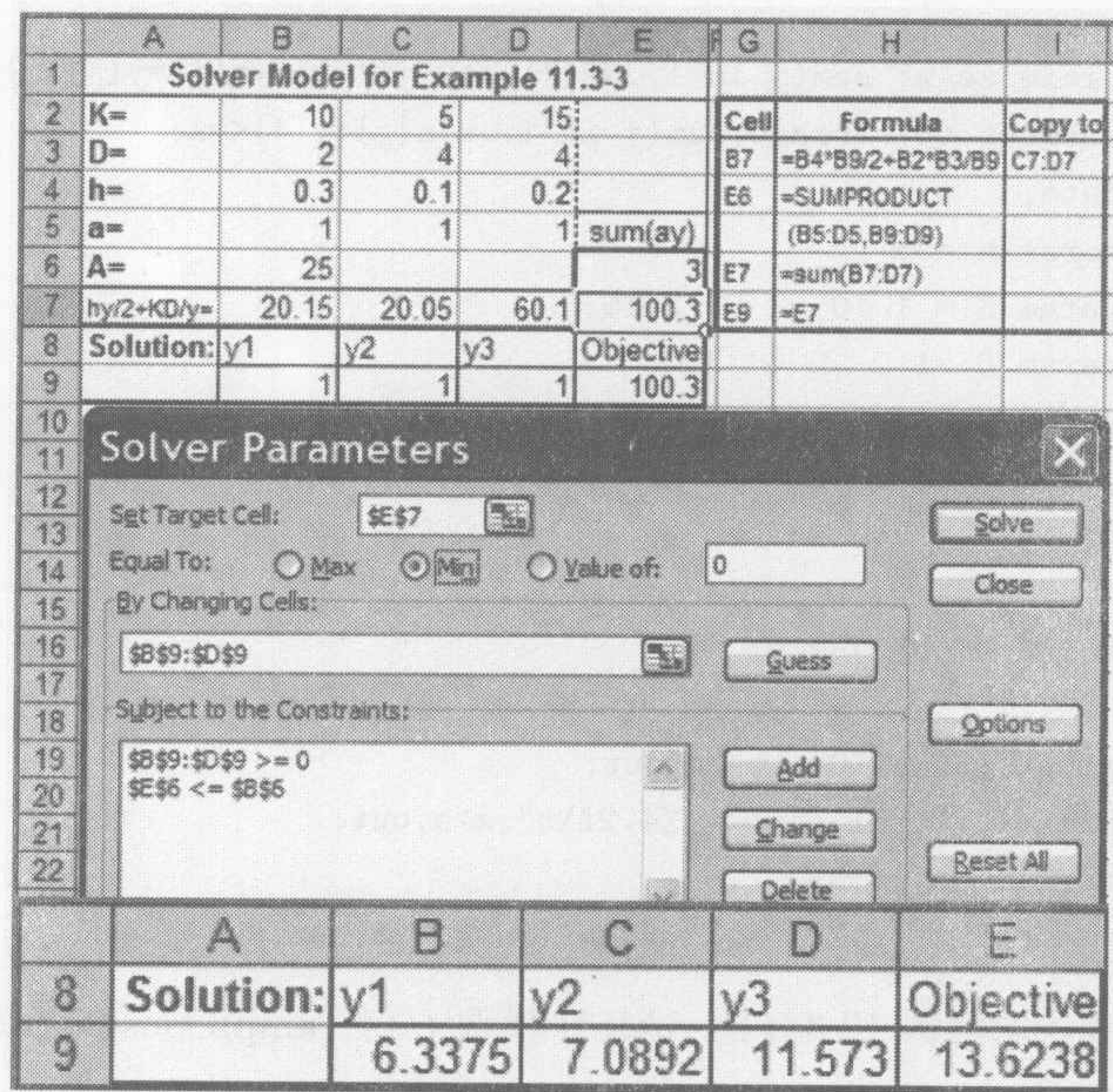


图 10.6 求解例 10.3-3 的 Excel 规划求解程序 (文件 solverConstrEOQ.xls)

AMPL 程序

图 10.7 是求解带有缺货限制的一般性多货品 EOQ 问题的 AMPL 非线性模型程序 (文件名 amplConstrEOQ.txt). 该模型采用与求解线性规划相同的准则, 但像 Excel 规划求解一样, AMPL 非线性模型显示出某些特性, 有可能得不到最优解. 特别是, 必须对变量指定“明智的”初始值. 在图 10.7 中, 定义语句

```
var y{1..n}>=0, :=10; #initial trial value = 10;
```

中的代码:=10, 是把初始值赋给所有变量. 假如你在现在的例子中使用初始值 1 的话, 迭代过程中就会导致被零除. 因此, 如在 Excel 规划求解一样, 可能需要用 $K_iD_i/(y_i+\Delta)$ 代替 K_iD_i/y_i, 其中 $\Delta > 0$ 是一个非常小量, 用在迭代中防止零除. 习题 10.3C 的第 1 题和第 4 题如果不用这一技巧, AMPL 就会求不出解来.

```
param n;
param K{1..n};
param D{1..n};
param h{1..n};
param a{1..n};
param A;

var y{1..n}>=0, :=10;   #initial trial value = 10

minimize z: sum{j in 1..n}(K[j]*D[j]/y[j]+h[j]*y[j]/2);
subject to storage:sum{j in 1..n}a[j]*y[j]<=A;
data;
param n:=3;
param K:= 1 10  2 5  3 15;
param D:=1 2  2 4  3 4;
param h:=1 .3 2 .1  3 .2;
param a:=1 1  2 1  3 1;
param A:=25;

solve;display z,y;

printf"SOLUTION:\n">a.out;
printf" Total cost = %4.2f\n",z>a.out;
for {i in 1..n}
    printf" y%1i = %4.2f\n",i,y[i]>a.out;
```

图 10.7 求解例 10.3-3 的 AMPL 模型 (文件 amplConstrEOQ.xls)

习题 10.3C[①]

1. 下列数据描述 5 个库存货品.

货品 i	K_i(\$)	D_i(每天库存数)	h_i(\$)	a_i(ft^2)
1	20	22	0.35	1.0
2	25	34	0.15	0.8
3	30	14	0.28	1.1
4	28	21	0.30	0.5
5	35	26	0.42	1.2
		可用的总储存面积 = 25 ft^2		

请求出最优订货量.

2. 解例 10.3-3 的模型, 假定所有货品平均库存的总量小于 25 件.

3. 对第 2 题假定, 唯一限制条件是投入库存资金的总量不超过 \$1 000. 货品 1, 2, 3 的单件采购费用分别为 \$100、\$55 和 \$100. 求最优解.

4. 下面是 4 种货品的库存数据:

① 你会发现, 文件 solverConstrEOQ.xls 和 amplConstrEOQ.txt 对求解本习题中的问题非常有用.

货品 i	K_i($)	D_i(每天库存数)	h_i($)
1	100	10	0.1
2	50	20	0.2
3	90	5	0.2
4	20	10	0.1

该公司想要确定这 4 种货品每种的经济订货量, 使得一年 365 天的订货总量不超过 150 件. 请将该问题用非线性规划表示, 并求出最优解.

10.4　动态 EOQ 模型

本节介绍的模型与 10.3 节相比有两点不同: (1) 在有限多个等长的周期期间, 周期性盘点库存水平; (2) 每个周期内的需求虽然是确定性的, 但从不同周期间的变化角度来看, 它是动态的.

出现动态的确定性需求的一种情况就是**原材料需求计划** (Material Requirement Planning, MRP). 我们举一个例子来阐述 MRP 的思想. 假设某种产品由两个模型 M_1 和 M_2 给出的下一年每季度的需求分别为 100 件和 150 件, 每个季度末进货. M_1 的生产提前时间为 2 个月, M_2 为 1 个月, M_1 和 M_2 的每件产品要用 2 件零部件 S, 而 S 的生产提前时间为 1 个月.

图 10.8 显示了 M_1 和 M_2 的生产计划安排, 这个计划从 3, 6, 9, 12 月底两个模型的季度需求开始 (实线箭头). 设 M_1 和 M_2 的提前时间为 2 个月和 1 个月, 虚线箭头表示每种产品批量的计划起点.

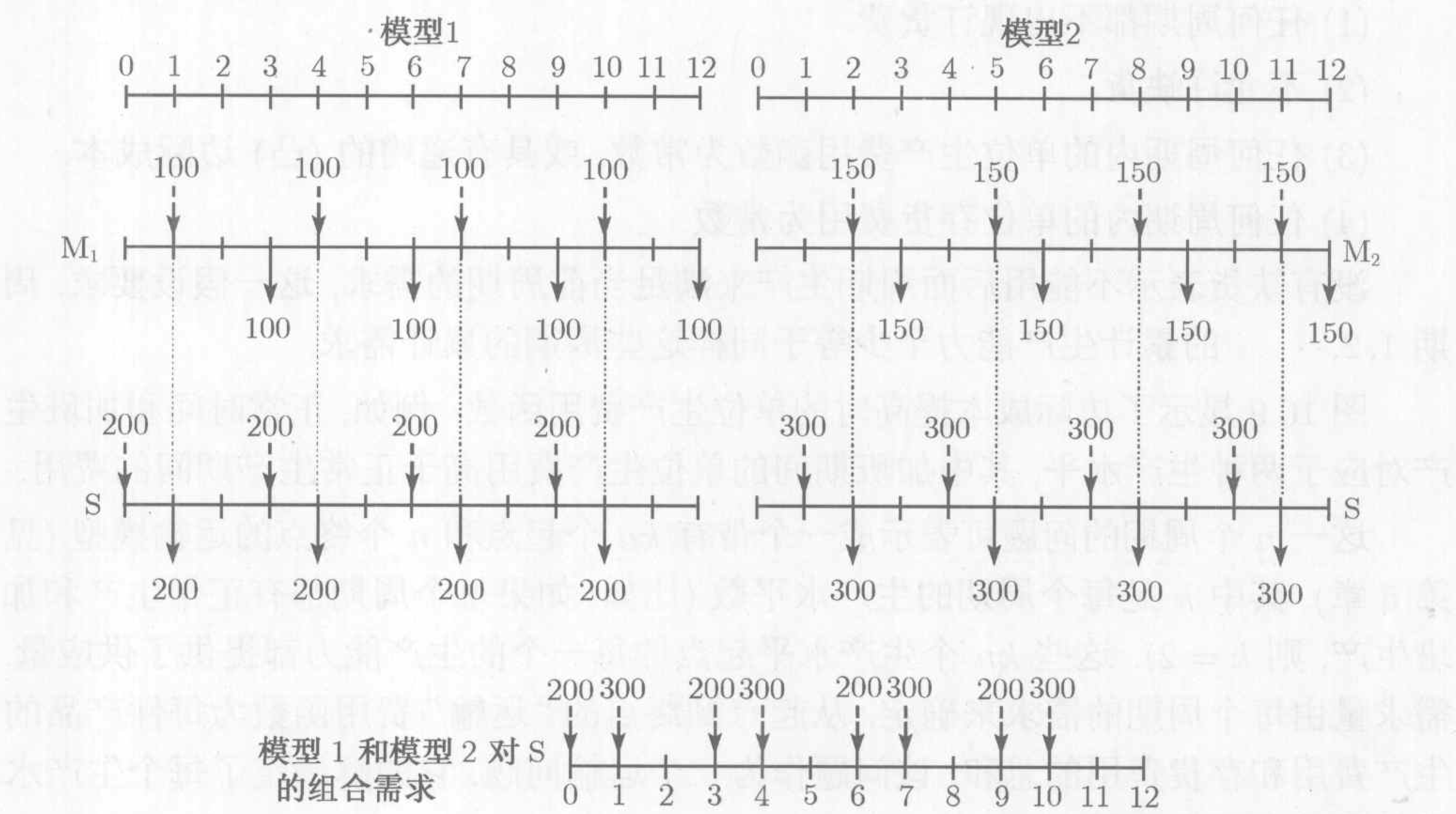


图 10.8　由 MRP 产生的动态需求的例子

为了按时启动这两个模型的生产, 零部件 S 的供货必须与虚线 M_1 和 M_2 箭头的时间一致. 这一信息用 S 图的实线箭头表示, 其中每个 M_1 或 M_2 单位产品对 S 的需求量为 2 件. 利用 1 个月的提前时间, S 图上的虚线箭头给出了对 S 的生产计划. 从这两个计划可以看出, 根据 M_1 和 M_2 可以确定对 S 的组合需求, 如图 10.8 最底下部分所示. 所得到的对 S 的可变(但已知的) 需求是常见的一类情况, 此时需要使用动态的 EOQ. 实质上的问题是, 给定对 S 的这种变化的需求, 为了减少总的库存成本, 每月初应该生产多少?

本节介绍两个模型. 第 1 个模型不带订货费, 第 2 个模型带有订货费. 这个看似"小小的"细节在模型复杂性方面会产生很大的不同.

习题 10.4A

1. 在图 10.8 中, 对下列每种情况, 求出对于零部件 S 的组合需求:

*(a) M_1 的提前时间只有 1 个周期.　　(b) M_1 的提前时间为 3 个周期.

10.4.1　不带订货费的模型

这一模型考虑划分成 n 段等长周期的计划. 每段周期的生产能力有限, 有若干个不同的生产水平 (例如, 正常班生产和加班生产代表两种生产水平). 当前周期可以生产出多于现行需求的产品, 以满足后面周期的需求, 在这种情况下, 必然需要存货费.

该模型的一般假设条件有:

(1) 任何周期都不出现订货费.

(2) 不允许缺货.

(3) 任何周期内的单位生产费用函数为常数, 或具有递增的 (凸) 边际成本.

(4) 任何周期内的单位存货费用为常数.

没有缺货表示不能用后面周期生产来满足当前周期的需求, 这一假设要求, 周期 $1, 2, \cdots, i$ 的累计生产能力至少等于同样这些周期的累计需求.

图 10.9 显示了边际成本提高时的单位生产费用函数. 例如, 正常时间和加班生产对应于两种生产水平, 其中加班期间的单位生产费用高于正常生产期间的费用.

这一 n 个周期的问题可表示成一个带有 kn 个起点和 n 个终点的运输模型 (见第 5 章), 其中 k 是每个周期的生产水平数 (比如, 如果每个周期都有正常生产和加班生产, 则 $k = 2$). 这些 kn 个生产水平起点的每一个的生产能力都提供了供应量. 需求量由每个周期的需求来确定. 从起点到终点的"运输"费用函数为每件产品的生产费用和存货费用的总和. 该问题作为一个运输问题, 它的解确定了每个生产水平的最小费用生产量.

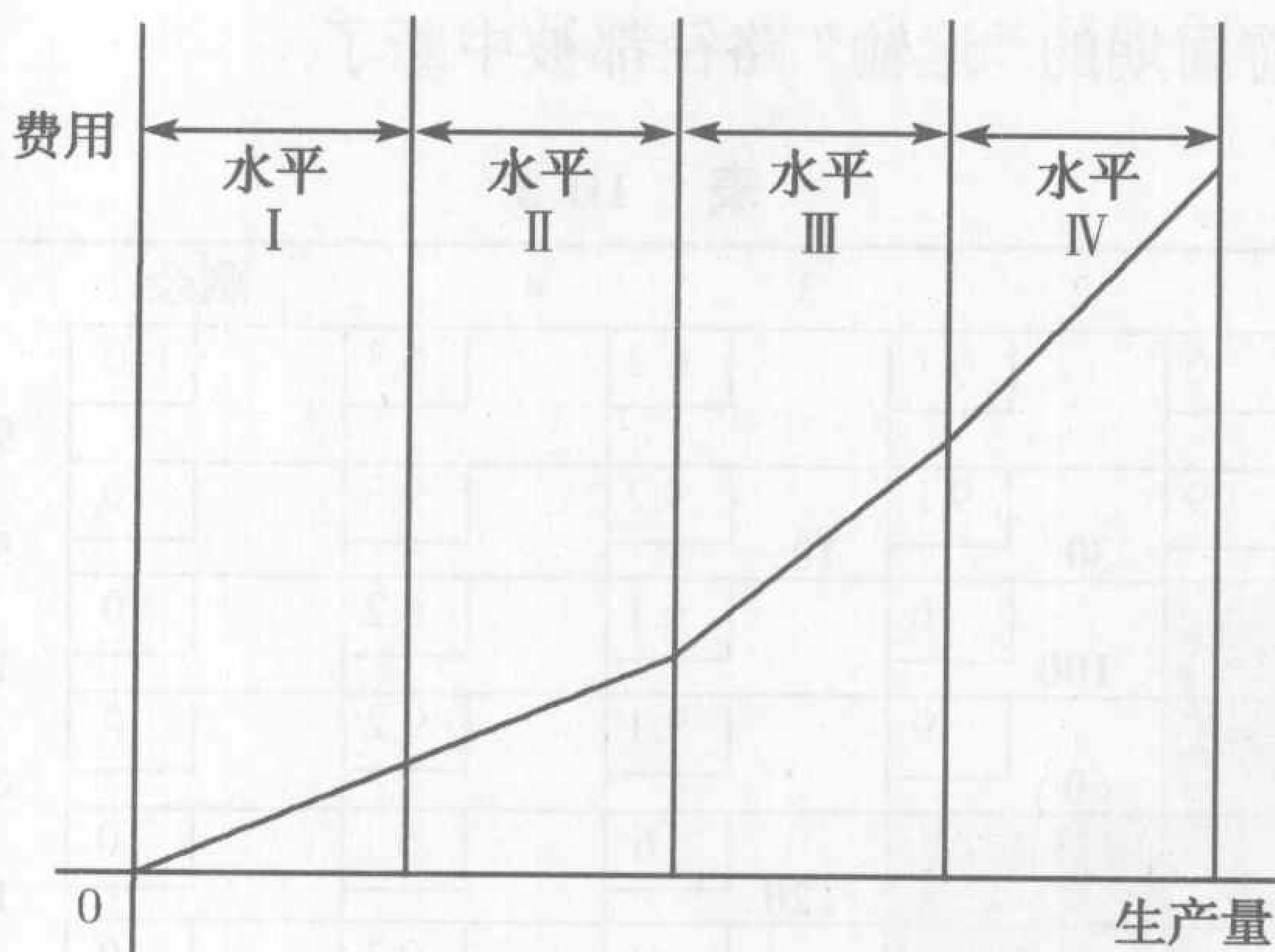


图 10.9 凸单位生产费用函数

不用第 5 章中我们熟悉的运输问题求解方法, 也可以求解这个运输模型, 新的求解算法利用了没有缺货和凸生产费用函数的假定.

例 10.4-1

Metalco 公司生产通风挡板, 可在 12 月到 3 月期间用在家庭壁炉里. 产品需求开始很低, 到冬季中期达到高峰, 到冬季末逐渐消失. 由于产品畅销, Metalco 公司可能会加班生产, 以满足需求. 下表给出了冬天这 4 个月的生产能力和需求情况.

月	生产能力		需求 (个)
	正常生产 (个)	加班生产 (个)	
1	90	50	100
2	100	60	190
3	120	80	210
4	110	70	160

任何生产周期的单位生产成本为: 正常生产期间 \$6, 加班期间 \$9. 每件产品每个月的储存费为 \$10.

为了在不允许缺货情况下保证模型有可行解, 任何月份的累计供货量 (生产能力) 必须至少等于相应的累计需求量, 如下表所示.

月	累计供应量	累计需求量
1	$90+50=140$	100
2	$140+100+60=300$	$100+190=290$
3	$300+120+80=500$	$290+210=500$
4	$500+110+70=680$	$500+160=660$

表 10.2 总结了模型的计算结果. 符号 R_i 和 O_i 代表周期 $i, i=1,2,3,4$ 期间的正常生产水平和加班生产水平. 因为第 4 周期的累计供应量超过了累计需求, 我们加上了一个虚拟剩余终点来平衡这个模型, 如表 10.2 中所示. 因为不允许缺货, 所

有从前面周期到当前周期的“运输”路径都被中断了.

表 10.2

	1	2	3	4	剩余	
R_1	6 / **90**	6.1	6.2	6.3	0	90
O_1	9 / **10**	9.1 / **30**	9.2 / **10**	9.3	0	50→40→10
R_2		6 / **100**	6.1	6.2	0	100
O_2		9 / **60**	9.1	9.2	0	60
R_3			6 / **120**	6.1	0	120
O_3			9 / **80**	9.1	0	80
R_4				6 / **110**	0	110
O_4				9 / **50**	0 / **20**	70→20
	100	190	210	160	20	
	↓	↓	↓	↓		
	10	90	90	50		
		↓	↓			
		30	10			

单位“运输”费用为已知的生产费用和储存费用. 例如, 从 R_1 到周期 1 的单位”费用等于仅有的单位生产成本 (= \$6), 而从 O_1 到周期 4 的单位费用等于单用加上从周期 1 到 4 的单位储存费用, 即 $\$9+(\$0.1+\$0.1+\$0.1)=\$9.30$. , 到剩余终点的单位费用等于零.

从第 1 列开始, 每次向着剩余列移动一列, 从头到尾走一遍就得到了最优解. 于每一列, 利用该列最低费用路径, 就能满足需求[①].

从第 1 列开始, 路径 $(R_1,1)$ 具有最低费用, 我们给它赋予最大可能值, 即 $\min\{90,100\}=90$ 件, 第 1 列还有 10 件没有得到满足. 第 1 列的下一个最低费路径是 $(O_1,1)$, 我们对它赋值 $\min\{50,10\}=10$. 现在周期 1 的需求得到了满足.

接下来看第 2 列. 这列的赋值按以下顺序: 100 件到 $(R_2,2)$, 60 件到 $(O_2,2)$, 30 件到 $(O_1,2)$. 这些赋值相应的单位“运输”费用为 \$6、\$9 和 \$9.10. 我们并没有用到 $(R_1,2)$, 它的单位费用为 \$6.10, 因为所有 R_1 的供货都赋给了周期 1.

按照同样的方法继续, 我们满足第 3 列的需求, 然后满足第 4 列的需求. 表 10.2 黑体所示的最优解如下:

相应的总费用为 $90\times\$6+10\times\$9+30\times\$9.10+100\times\$6+60\times\$9+10\times\$9.20+120\times\$6+80\times\$9+110\times\$6+50\times\$9=\$4\,685$.

① 这一算法的最优性证明请见 S.M., Johnson, “Sequential Production Planning over Time at Minimum Cost,” *Management Science*, Vol. 3, pp. 435-437, 1957.

周期	生产计划安排
正常生产 1	周期 1 生产 90 件
加班生产 1	生产 50 件：周期 1 生产 10 件，周期 2 生产 30 件，周期 3 生产 10 件
正常生产 2	周期 2 生产 100 件
加班生产 2	周期 2 生产 60 件
正常生产 3	周期 3 生产 120 件
加班生产 3	周期 3 生产 80 件
正常生产 4	周期 4 生产 110 件
加班生产 4	周期 4 生产 50 件，剩余生产能力 20 件

习题 10.4B

1. 求解例 10.4-1，设单位生产费用和储存费用 (单位：美元) 如下表.

周期 i	正常生产 单位产品的成本	加班生产 单位产品的成本	到周期 $i+1$ 的 单位储存费用
1	5.00	7.50	0.10
2	3.00	4.50	0.15
3	4.00	6.00	0.12
4	1.00	1.50	0.20

2. 根据下表数据生产某产品以满足 4 个周期的已知需求：

产量范围 (件)	各周期的单位生产费用 ($)			
	1	2	3	4
1~3	1	2	2	3
4~11	1	4	5	4
12~15	2	4	7	5
16~25	5	6	10	7
下一周期单位储存费用 ($)	0.30	0.35	0.20	0.25
总需求 (件)	11	4	17	29

(a) 找出最优解，给出每个周期要生产的件数.

(b) 假定周期 4 另外需要 10 件产品，则应在哪里生产？

***3.** 某产品未来 5 个周期的需求可以按照正常生产、加班生产或转包生产来实现. 只有当加班生产情况下才采用转包生产方式. 下表给出了该情形下的供应、需求和费用数据.

周期	生产能力 (件)			需求
	正常生产	加班生产	转包生产	
1	100	50	30	153
2	40	60	80	200
3	90	80	70	150
4	60	50	20	200
5	70	50	100	203

每个周期内这 3 种不同生产水平的单位生产成本分别是 \$4、\$6 和 \$7. 每个周期的单位储存费用为 \$0.50. 请求出最优解.

10.4.2 带有订货费的模型

本问题不允许缺货, 且每次开始一个新的生产批量时, 都需要一笔订货费. 本节介绍两种求解方法: 一种是纯粹的动态规划方法, 另一种是启发式算法.

图 10.10 表示了这种库存情形. 对周期 $i, i = 1, 2, \cdots, n$, 图上的符号定义如下:

$$z_i = \text{订货量}$$

$$D_i = \text{周期 } i \text{ 的需求}$$

$$x_i = \text{周期 } i \text{ 初始库存}$$

费用项定义为

$$K_i = \text{周期 } i \text{ 的订货费}$$

$$h_i = \text{从周期 } i \text{ 到 } i+1 \text{ 的单位库存费用}$$

周期 i 相应的生产费用函数是

$$C_i(z_i) = \begin{cases} 0, & z_i = 0 \\ K_i + c_i(z_i), & z_i > 0 \end{cases}$$

函数 $c_i(z_i)$ 为 z_i 的边际生产费用函数.

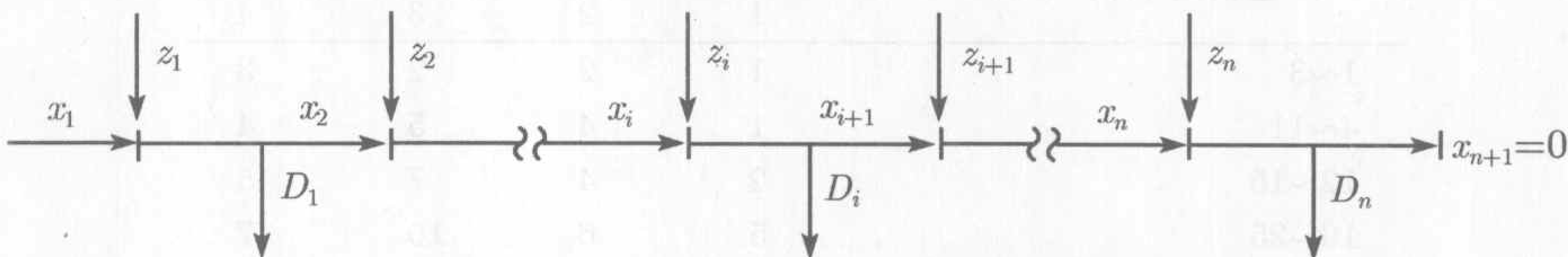


图 10.10 带有订货费的动态库存模型要素

一般动态规划算法 在没有缺货情况下, 该库存模型求所有 n 个周期生产费用与储存费用之和的极小值. 为了简单起见, 我们假定阶段 (周期)i 的储存费用是根据周期末库存来计算的, 定义为

$$x_{i+1} = x_i + z_i - D_i$$

对于前向递归公式, 定义 x_{i+1} 为阶段 i 的状态, 即周期末的库存水平. 如图 10.10 所示,

$$0 \leqslant x_{i+1} \leqslant D_{i+1} + \cdots + D_n$$

从这个不等式可以看出, 在极端情况下, 剩余库存量 x_{i+1} 能够满足所有剩余周期的需求.

令 $f_i(x_{i+1})$ 为已知周期末库存量 x_{i+1} 时周期 $1, 2, \cdots, i$ 的最小库存费用, 则前向递归公式为

$$f_1(x_2) = \min_{z_1 = D_1 + x_2 - x_1} \{C_1(z_1) + h_1 x_2\}$$

$$f_i(x_{i+1}) = \min_{0 \leqslant z_i \leqslant D_i + x_{i+1}} \{C_i(z_i) + h_i x_{i+1} + f_{i-1}(x_{i+1} + D_i - z_i)\}, i = 2, 3, \cdots, n$$

注意对于周期 1, z_1 必须严格等于 $D_1 + x_2 - x_1$. 对于 $i > 1, z_i$ 最小可以为 0, 因为 D_i 可以从前面的周期 $1, 2, \cdots, i-1$ 的生产中得到满足.

例 10.4-2

下表给出了某 3 个周期库存情形的数据.

周期 i	需求 D_i(件)	订货费 K_i(\$)	储存费 h_i(\$)
1	3	3	1
2	2	7	3
3	4	6	2

需求是离散的, 初始库存为 $x_1 = 1$ 件. 头 3 件产品的单位生产费用为 \$10, 其后每件产品为 \$20, 翻译为数学语言即是

$$c_i(z_i) = \begin{cases} 10z_i, & 0 \leqslant z_i \leqslant 3 \\ 30 + 20(z_i - 3), & z_i \geqslant 4 \end{cases}$$

确定最优库存策略.

周期 1 $\quad D_1 = 3, 0 \leqslant x_2 \leqslant 2 + 4 = 6, z_1 = x_2 + D_1 - x_1 = x_2 + 2$

		$C_1(z_1) + h_1 x_2$							最优解	
		$z_1 = 2$	3	4	5	6	7	8		
x_2	$h_1 x_2$	$C_1(z_1) = 23$	33	53	73	93	113	133	$f_1(x_2)$	z_1^*
0	0	23							23	2
1	1		34						34	3
2	2			55					55	4
3	3				76				76	5
4	4					97			97	6
5	5						118		118	7
6	6							139	139	8

注意, 因为 $x_1 = 1$, 所以 z_1 的最小值为 $D_1 - x_1 = 3 - 1 = 2$.

周期 2 $\quad D_2 = 2, 0 \leqslant x_3 \leqslant 4, 0 \leqslant z_2 \leqslant D_2 + x_3 = x_3 + 2$

		$C_2(z_2)+h_2x_3+f_1(x_3+D_2-z_2)$							最优解	
		$z_2=0$	1	2	3	4	5	6		
x_3	h_2x_3	$C_2(z_2)=0$	17	27	37	57	77	97	$f_2(x_3)$	z_2^*
0	0	0+55 =55	17+34 =51	27+23 =50					50	2
1	3	3+76 =79	20+55 =75	30+34 =64	40+23 =63				63	3
2	6	6+97 =103	23+76 =99	33+55 =88	43+34 =77	63+23 =86			77	3
3	9	9+118 =127	26+97 =123	36+76 =112	46+55 =101	66+34 =100	86+23 =109		100	4
4	12	12+139 =151	29+118 =147	39+97 =136	49+76 =125	69+55 =124	89+34 =123	109+23 =132	123	5

周期 3 $D_3=4, x_4=0, 0\leqslant z_3\leqslant D_3+x_4=4$

		$C_3(z_3)+h_3x_4+f_2(x_4+D_3-z_3)$					最优解	
		$z_3=0$	1	2	3	4		
x_4	h_3x_4	$C_3(z_3)=0$	16	26	36	56	$f_3(x_4)$	z_3^*
0	0	0+123 =123	16+100 =116	26+77 =103	36+63 =99	56+50 =106	99	3

最优解的解读方式为

$$(x_4=0)\to\boxed{z_3=3}\to(x_3=0+4-3=1)\to\boxed{z_2=3}$$
$$\to(x_2=1+2-3=0)\to\boxed{z_1=2}.$$

因此, 最终的最优解为 $z_1^*=2, z_2^*=3, z_3^*=3$, 总费用 \$99.

Excel 程序

Excel 程序 excelDPInv.xls 可用来求解多达 10 个周期的一般动态规划库存问题. 该电子表格的设计和 9.3.1 节中给出的 excelKanpsack.xls 程序非常类似, 每次计算一个阶段, 需要用户输入把各阶段连接在一起.

图 10.11 显示了 excelDPInv.xls 用来求解例 10.4-2 的情形. 输入每个阶段的数据, 从阶段 1 开始计算, 并注意到, 费用函数 $c_i(z_i)$ 在第 3 行输入: (G3 = 10, H3 = 20, I3 = 3), 意思是头 3 件产品的单位费用为 \$10, 其后每件产品的单位费用为 \$20. 还要注意到对 D_1 输入的量必须为初始库存已经被扣除后的净量 ($=3-x_1=3-1=2$). 此外, 你还需要建立变量 z_1 的可行值. 电子表格会自动检查你输入的值是否正确, 并在第 6 行给出自解释性的消息: yes、no 或 delete.

输入所有数据以后, 在 S 列和 T 列会给出该阶段 f_i 和 z_i 的最优解值. 接下来, 在电子表格最优解结果区域会对阶段 1 的解 (x_1, f_1, z_1) 做一个永久性纪录, 如图 10.11 所示. 这要求复制 D9:D15 和 S9:T15 单元格, 并用 paste special+values 菜

单项对它们进行粘贴 (你可能需要回顾 9.3.1 节的 excelKnapsack.xls 程序以建立永久性纪录.)

Period 1:

	A	B	C	D	E	F	G	H	I	J	K	S	T	U	V	W	X	Y	Z
1		General (Forward) Dynamic Programming Inventory Model																	
2	I	Number of periods, N=			3	Current period=		1						Optimum solution					
3	N	K1=	3	h1=	1	c1(z1)=	10	20	3					Summary					
4	P	Period	1	2	3									x	f	z	x	f	z
5	U	D(1 to 3)=	2	2	4									Period 1					
6	T	Are z1 values correct?			yes	yes	yes	yes	yes	yes	yes	Optimum		0	23	2			
7		Period 0		z1=	2	3	4	5	6	7	8	Period1		1	34	3			
8		f0		C1(z1)=	23	33	53	73	93	113	133	f1	z1	2	55	4			
9	S		x2=	0	23	1111111	1111111	1111111	1111111	1111111	1111111	23	2	3	76	5			
10	T		x2=	1	1111111	34	1111111	1111111	1111111	1111111	1111111	34	3	4	97	6			
11	A		x2=	2	1111111	1111111	55	1111111	1111111	1111111	1111111	55	4	5	118	7			
12	G		x2=	3	1111111	1111111	1111111	76	1111111	1111111	1111111	76	5	6	139	8			
13	E		x2=	4	1111111	1111111	1111111	1111111	97	1111111	1111111	97	6						
14			x2=	5	1111111	1111111	1111111	1111111	1111111	118	1111111	118	7						
15	C		x2=	6	1111111	1111111	1111111	1111111	1111111	1111111	139	139	8						

Period 2:

	A	B	C	D	E	F	G	H	I	J	K	S	T	U	V	W	X	Y	Z
1		General (Forward) Dynamic Programming Inventory Model																	
2	I	Number of periods, N-			3	Current period-		2						Optimum solution					
3	N	K2=	7	h2=	3	c2(z2)=	10	20	3					Summary					
4	P	Period	1	2	3									x	f	z	x	f	z
5	U	D(1 to 3)=	2	2	4									Period 1			Period 2		
6	T	Are z2 values correct?			yes	yes	yes	yes	yes	yes	yes	Optimum		0	23	2	0	50	2
7		Period 1		z2-	0	1	2	3	4	5	6	Period2		1	34	3	1	63	3
8		f1		C2(z2)=	0	17	27	37	57	77	97	f2	z2	2	55	4	2	77	3
9	S	23	x3=	0	55	51	50	1111111	1111111	1111111	1111111	50	2	3	76	5	3	100	4
10	T	34	x3=	1	79	75	64	63	1111111	1111111	1111111	63	3	4	97	6	4	123	5
11	A	55	x3=	2	103	99	88	77	86	1111111	1111111	77	3	5	118	7			
12	G	76	x3=	3	127	123	112	101	100	109	1111111	100	4	6	139	8			
13	E	97	x3=	4	151	147	136	125	124	123	132	123	5						
14		118																	
15	C	139																	

Period 3:

	A	B	C	D	E	F	G	H	I	J	K	S	T	U	V	W	X	Y	Z
1		General (Forward) Dynamic Programming Inventory Model																	
2	I	Number of periods, N=			3	Current period=		3						Optimum solution					
3	N	K3=	6	h3=	2	c3(z3)=	10	20	3					Summary					
4	P	Period	1	2	3									x	f	z	x	f	z
5	U	D(1 to 3)=	2	2	4									Period 1			Period 2		
6	T	Are z3 values correct?			yes	yes	yes	yes	yes			Optimum		0	23	2	0	50	2
7		Period 2		z3=	0	1	2	3	4			Period3		1	34	3	1	63	3
8		f2		C3(z3)=	0	16	26	36	56			f3	z3	2	55	4	2	77	3
9	S	50	x4=	0	123	116	103	99	106			99	3	3	76	5	3	100	4
10	T	63												4	97	6	4	123	5
11	A	77												5	118	7	Period 3		
12	G	100												6	139	8	0	99	3
13	E	123																	

图 10.11 例 10.4-2 的 Excel 动态规划解 (文件 excelDPInv.xls)

下一步, 为了准备阶段 2 的计算, 从永久性纪录中复制 f_1, 粘贴到 B 列, 如图 10.11 所示. 接下来的工作就是更新阶段 2 的输入数据. 对阶段 3 重复上述过程.

习题 10.4C

***1.** 考虑例 10.4-2.

(a) 让 $x_4 > 0$ 是否有意义?

(b) 对下列两种情况的每一种, 确定 $z_1, z_2, z_3, x_1, x_2, x_3$ 的可行范围 (如图 10.10 那样来

表示每种情况可能非常有用).

(i) $x_1 = 4$, 所有其他数据不变.

(ii) $x_1 = 0, D_1 = 5, D_2 = 3, D_3 = 4$

2. ***(a)** 找出下面 4 个周期库存模型的最优解.

周期 i	需求 D_i(件)	订货费 K_i($)	储存费 h_i($)
1	5	5	1
2	2	7	1
3	3	9	1
4	3	7	1

前 6 件产品的单位生产费用为 $1, 其后每件产品为 $2.

(b) 利用 excelDPInv.xls 检验你的计算结果.

3. 假设库存–储存费用按各阶段的平均库存计算, 请导出相应的前向递归公式.

4. 给出该模型的后向递归公式, 然后用它来解例 10.4-2.

5. 建立该模型的后向递归公式, 假定库存–储存费用按各阶段的平均库存计算.

带有常数或递减边际成本的动态规划算法 上面给出的一般 DP 模型可用于任何费用函数, 但是由算法的性质可知, 阶段 i 的状态 x_i 和决策方案 z_i 每次增加 1. 这就意味着对于需求数量较大的问题, 每个阶段的计算表会非常大, 因此计算上非常繁琐.

一般动态规划的一种特殊情况有望减少计算量. 在这种情况下, 单位生产费用和单位储存费用都分别是生产量和库存水平的非增(凹) 函数. 这一情况往往发生在当单位费用函数为常数, 或者允许订货量折扣的情形下.

在给定的条件下, 可以证明①:

(1) 给定零初始库存 ($x_1 = 0$), 满足任何周期 i 的需求的最优方案可以通过以下两者之一, 要么通过新的生产, 要么通过增加库存. 但不能同时通过这两者, 即 $z_i x_i = 0$. (对于正的初始库存情况, $x_1 > 0$, 这个量可以从后续周期的需求中扣除, 直到耗尽.)

(2) 周期 i 的最优生产量 z_i 要么必须是零, 要么必须刚好满足一个或几个后续周期的需求.

例 10.4-3

某公司 4 个周期的库存模型的数据如下:

周期 i	需求 D_i(件)	订货费 K_i($)
1	76	98
2	26	114
3	90	185
4	67	70

① 见 H.Wagner and T. Whitin, "Dynamic Version of the Economic Lot Size Model", *Management Science*, Vol. 5. pp. 89-96, 1958. 这一最优性证明要求限制条件假设, 即对所有周期设费用函数为常数或一致, 后来 A. Veinott Jr. 放松了这一假设, 允许有不同的凹费用函数.

初始库存为 $x_1 = 15$ 件, 单位生产费用为 \$2, 每个周期的单位储存费用对所有周期而言都是 \$1(为简单起见, 我们假设单位生产和储存费用对所有周期不变).

使用前面给出的后向递归算法求最优解, 不同的是, 我们让 x_{i+1} 和 z_i 的值"一次性付清", 而不是每次增加 1. 因为 $x_1 = 15$, 所以第 1 周期的需求调整为 $76 - 15 = 61$ 件.

周期 1 $D_1 = 61$

		$C_1(z_1)+h_1x_2$				最优解	
		$z_1=61$	87	177	244		
x_2	h_1x_2	$C_1(z_1)=220$	272	452	586	$f_1(x_2)$	z_1^*
0	0	220				220	61
26	26		298			298	87
116	116			568		568	177
183	183				769	769	244
在周期 1	订货给	1	1,2	1,2,3	1,2,3,4		

周期 2 $D_2 = 26$

		$C_2(z_2)+h_2x_3+f_1(x_3+D_2-z_2)$				最优解	
		$z_2=0$	26	116	183		
x_3	h_2x_3	$C_2(z_2)=0$	166	346	480	$f_2(x_3)$	z_2^*
0	0	$0+298=298$	$166+220=386$			298	0
90	90	$90+568=658$		$436+220=656$		656	116
157	157	$157+769=926$			$637+220=857$	857	183
在周期 2	订货给	—	2	2,3	2,3,4		

周期 3 $D_3 = 90$

		$C_3(z_3)+h_3x_4+f_2(x_4+D_3-z_3)$			最优解	
		$z_3=0$	90	157		
x_4	h_3x_4	$C_3(z_3)=0$	365	499	$f_3(x_4)$	z_3^*
0	0	0+656=656	365+298=663		656	0
67	67	67+857=924		566+298=864	864	157
在周期 3	订货给	—	3	3,4		

周期 4 $D_4 = 67$

		$C_4(z_4)+h_4x_5+f_3(x_5+D_4-z_4)$		最优解	
		$z_4=0$	67		
x_5	h_4x_5	$C_4(z_4)=0$	204	$f_4(x_5)$	$z_4{}^*$
0	0	0+864=864	204+656=860	860	67
在周期 4	订货给	—	4		

上表得出的最优策略如下:

$$(x_5=0)\to \boxed{z_4=67}\to(x_4=0)\to\boxed{z_3=0}$$
$$\to(x_3=90)\to\boxed{z_2=116}\to(x_2=0)\to\boxed{z_1=61}.$$

我们得到 $z_1^*=61, z_2^*=116, z_3^*=0, z_4^*=67$, 总费用为 \$860.

Excel 程序

程序 excelWagnerWhitin.xls 类似于一般模型的 excelDPInv.xls, 唯一不同的是对状态 x 和决策方案 z 采用了一次付清. 还有, 为了简单起见, 这一新的电子表格程序不允许批量折扣. 图 10.12 对例 10.4-3 给出了周期 1 的计算结果. 该程序最多可以有 10 个周期. 记住, 在建立输出解结果 (Q:V 列) 时使用 paste special+values 菜单命令.

Period 1:

	A	B	C	D	E	F	G	H	I	O	P	Q	R	S
1		Wagner-Whitin (Forward) Dynamic Programming Inventory Model												
2		Number of periods, N=			4		Current period=		1					
3	I	Period	1	2	3	4								
4	N	c(1 to 4) =	2	2	2	2						Optimum Solution		
5	P	K(1 to 4) =	98	114	185	70							Summary	
6	U	h(1 to 4) =	1	1	1	3						x	f	z
7	T	D(1 to 4) =	61	26	90	67				Current			period 1	
8		Are z1 values correct?			yes	yes	yes	yes		optimum		0	220	61
9		Period 0		z1=	61	87	177	244		Period 1		26	298	87
10	4	f 0		C1(z1)=	220	272	452	586		f1	z1	116	568	177
11	S		x2=	0	220	111111	1111111	1111111		220	61	183	769	244
12	T		x2=	26	1111111	298	1111111	1111111		298	87			
13	A		x2=	116	1111111	1111111	568	1111111		568	177			
14	G		x2=	183	1111111	1111111	1111111	769		769	244			

图 10.12　用于例 10.4-3 中周期 1 的 Wagner-Whitin Excel 动态规划模型 (文件 excelWagnerWhitin.xls)

习题 10.4D

***1.** 设初始库存为 80 件, 求解例 10.4-3, 并利用 excelWagnerWhitin.xls 检查计算结果.

2. 解下面 10 个周期的确定性库存模型, 设初始库存为 50 件.

周期 i	需求 D_i(件)	单位生产成本 (\$)	单位储存费 (\$)	订货费 (\$)
1	150	6	1	100
2	100	6	1	100
3	20	4	2	100
4	40	4	1	200
5	70	6	2	200
6	90	8	3	200
7	130	4	1	300
8	180	4	4	300
9	140	2	2	300
10	50	6	1	300

3. 求出下列 5 个周期模型的最优库存策略. 所有周期的单位生产费用为 \$10, 每周期的单位储存费用为 \$1.

周期 i	需求 D_i(件)	订货费 K_i(\$)
1	50	80
2	70	70
3	100	60
4	30	80
5	60	60

4. 求出下面 6 个周期模型的最优库存策略. 所有周期的单位生产费用为 \$2.

周期 i	D_i(件)	K_i(\$)	h_i(\$)
1	10	20	1
2	15	17	1
3	7	10	1
4	20	18	3
5	13	5	1
6	25	50	1

Silver-Meal 启发式算法 这一启发式算法只适用于单位生产费用为常数而且对所有周期都一致的库存问题. 正因为如此, 算法只对订货费和储存费进行平衡.

这一启发式算法把用当前周期的需求来满足后续未来需求的那些周期找出来, 目标是使得每个周期相应的订货费和储存费用达到最小.

假设在周期 i 时生产产品, 以满足周期 $i, i+1, \cdots, t$ 的需求, $i \leqslant t$, 且定义 $TC(i,t)$ 为同样这些周期的订货费和储存费. 我们用动态规划模型的符号, 可将其表示为下面的数学公式

$$TC(i,t) = \begin{cases} K_i, & t = i \\ K_i + h_i D_{i+1} + (h_i + h_{i+1}) D_{i+2} + \cdots + \left(\sum_{k=i}^{t-1} h_k\right) D_t, & t > i \end{cases}$$

接下来定义 $TCU(i,t)$ 为每周期相应的费用, 即

$$TCU(i,t) = \frac{TC(i,t)}{t-i+1}$$

因此对当前周期 i, 该启发式算法求使得 $TCU(i,t)$ 达到极小的 t^*.

函数 $TC(i,t)$ 可按以下公式递归计算:

$$TC(i,i) = K_i$$

$$TC(i,t) = TC(i,t-1) + \left(\sum_{k=i}^{t-1} h_k\right) D_t, \quad t = i+1, i+2, \cdots, n$$

第 0 步 设定 $i = 1$.

第 1 步 求出满足下列两个条件的局部极小值 t^*:

$$TCU(i, t^*-1) \geqslant TCU(i, t^*)$$

$$TCU(i, t^*+1) \geqslant TCU(i, t^*)$$

如果满足上述条件, 则启发式算法要求在周期 i 为周期 $i, i+1, \cdots, t^*$ 订货 $(D_i + D_{i+1} + \cdots + D_{t^*})$.

第 2 步 设 $i=t^*+1$, 如果 $i>n$, 停止, 整个计划期间已经算完; 否则, 转向第 1 步.

例 10.4-4

针对下面 6 个周期的库存情形, 求出最优库存策略:

周期 i	D_i(件)	K_i($)	h_i($)
1	10	20	1
2	15	17	1
3	7	10	1
4	20	18	3
5	13	5	1
6	25	50	1

所有周期的单位生产成本为 $2.

迭代 1$(i=1, K_1=\$20)$ 对 t 递归计算函数 $TC(1,t)$. 例如, 给定 $TC(1,1)=\$20$, 则 $TC(1,2)=TC(1,1)+h_1D_2=20+1\times 15=\$35$.

周期 t	D_i	$TC(1,t)$	$TCU(1,t)$
1	10	$20	$\frac{20}{1}=\$20.00$
2	15	$20+1\times 15=\$35$	$\frac{35}{2}=\$17.50$
3	7	$35+(1+1)\times 7=\$94$	$\frac{49}{3}=\boxed{\$16.33}$
4	20	$49+(1+1+1)\times 20=\$109$	$\frac{109}{4}=\$27.25$

当 $t^*=3$ 时达到局部极小值, 因此要求在周期 1 里为周期 1 到周期 3 订货 $10+15+7=32$ 件. 设置 $i=t^*+1=3+1=4$

迭代 2$(i=4, K_4=\$18)$

周期 t	D_i	$TC(4,t)$	$TCU(4,t)$
4	20	$18	$\frac{18}{1}=\boxed{\$18.00}$
5	13	$18+3\times 13=\$57$	$\frac{57}{2}=\$28.50$

上述计算得到 $t^*=4$, 要求在周期 4 里为周期 4 订货 20 件. 设置 $i=4+1=5$.

迭代 3$(i=5, K_5=\$5)$

周期 t	D_i	$TC(5,t)$	$TCU(5,t)$
5	13	$5	$\frac{5}{1}=\boxed{\$5}$
6	25	$5+1\times 25=\$30$	$\frac{30}{2}=\$15$

当 $t^* = 5$ 时达到最小值, 要求在周期 5 里为周期 5 订货 13 件. 接下来, 我们设置 $i = 5 + 1 = 6$. 但由于 $i = 6$ 是计划期的最后一个周期, 因此我们必须在周期 6 里为周期 6 订货 25 件.

评注 下表对启发式算法与动态规划算法的解作了比较. 我们在动态规划模型中去掉了单位生产费用, 因为在启发式计算里没有这一项.

	启发式算法		动态规划算法	
周期	生产件数	费用 ($)	生产件数	费用 ($)
1	32	49	10	20
2	0	0	22	24
3	0	0	0	0
4	20	18	20	18
5	13	5	38	30
6	25	50	0	0
合计	90	122	90	92

由启发式算法给出的生产计划安排比动态规划解的费用要高 32%($122 比 $92). 启发式算法的表现“不够好”, 可能是由于问题中的数据引起的, 在这个问题中, 周期 5 和周期 6 的订货费变化非常大. 尽管如此, 本例表明, 启发式算法并没有“向前寻找”更好的生产计划的能力. 例如, 在周期 5 里为周期 5 和周期 6 订货 (而不是为每个周期分别订货) 能节省 $25, 这就能使得总的启发式算法解的费用降低到 $97.

Excel 程序

Excel 程序 excelSilverMeal.xls 用来执行所有的迭代计算, 并给出最终解. 该程序要求从输入计算所需要的数据开始, 包括所有周期的 N, K, h, D (这些要输入的数据在电子表格中用青绿色突出显示), 然后用户需要人工启动每次迭代, 直到完成所有周期的计算.

图 10.13 展示了 Excel 启发式算法程序对例 10.4-4 的计算. 在单元格 J11 中输入值 1 开始初始迭代, 表示在周期 1 开始第 1 次迭代, 这时电子表格程序会生成对应周期数 N(本例中 =6) 的多个行, 这个周期数会按照升序列在单元格 K11:K16 中. 现在我们查看第 P 列 (用青绿色标出), 并找到对应于局部极小值的周期, 其 $t = 3$, $TCU = \$16.33$. 这意味着下一次迭代将从周期 4 开始. 跳过一个空行, 在 J18 中输入值 4, 开始进行迭代 2 的计算, 这将显示局部极小值取在周期 4($TCU = \$18$), 并会指示出迭代 3 将从周期 5 开始. 重复这个过程, 从 J22 输入 5, 迭代 3 的局部极小值将取在节点 5. 接下来, 在 J25 中输入值 6 将会终止本问题的迭代. 随着每次迭代的完成, 电子表格程序将会自动显示相应的最优策略以及总费用, 如图 10.13 所示.

	J	K	L	M	N	O	P	Q–V
1	Silver-Meal Heuristic Inventory Model							
2	Input data:							
3	Number of periods, N =	6	<<Maximum 14 periods					
4	Period t=	1	2	3	4	5	6	
5	Setup cost, Kt =	20	17	10	18	5	50	
6	Holding cost, ht =	1	1	1	3	1	1	
7	Demand, Dt =	10	15	7	20	13	25	
8								
9	Solution complete	Model calculations:						Optimum solution (Total cost = $122.00):
10	Start Iteration at Period	Period	D_t	$\sum D_t$	$\sum h_t$	TC	TCU	
11	1	1	10	10	0.00	20.00	20.00	
12		2	15	25	1.00	35.00	17.50	
13		3	7	32	2.00	49.00	16.33	
14		4	20	52	3.00	109.00	27.25	
15		5	13	65	6.00	187.00	37.40	
16		6	25	90	7.00	362.00	60.33	Order 32 in period 1 for periods 1 to 3, cost = $49.00
17								
18	4	4	20	20	0.00	18.00	18.00	
19		5	13	33	3.00	57.00	28.50	
20		6	25	58	4.00	157.00	52.33	Order 20 in period 4 for periods 4 to 4, cost = $18.00
21								
22	5	5	13	13	0.00	5.00	5.00	
23		6	25	38	1.00	30.00	15.00	Order 13 in period 5 for periods 5 to 5, cost = $5.00
24								
25	6	6	25	25	0.00	50.00	50.00	Order 25 in period 6 for periods 6 to 6, cost = $50.00
26								

图 10.13 例 10.4-4 运用 Silver-Meal 启发式算法的 Excel 解 (文件 ExcelSilverMeal.xls)

习题 10.4E

*1. 对鱼杆的需求量在 12 月份最小, 4 月份达到最大. Fishing Hole 公司预测, 12 月份的需求量为 50 支, 随后每个月增加 10 支, 到 4 月份达到 90 支. 其后每月需求量减少 5 支. 除了从 2 月到 4 月的高峰需求月份以外, 一个批量的订货费为 $250, 而高峰月份的订货费增加到 $300. 每支鱼杆的生产费用大约是 $15, 全年不变, 而每支鱼杆每月的储存费用为 $1. Fishing Hole 公司正在制定下一年度 (1 月到 12 月) 的生产计划, 该公司该如何安排它的生产能力呢?

2. 一家小出版公司再版一本小说, 以满足未来 12 个月的需求. 对以后各月份的需求预测分别为 100, 120, 50, 70, 90, 105, 115, 95, 80, 100, 110 本. 再版该书的开机费用为 $200.00, 每本书每月的库存费用为 $1.20. 确定最优的再版计划安排.

参 考 文 献

Bishop, J. "Experience with a Successful System for Forecasting and Inventory Control," *Operations Research*, Vol.22, No.6, pp.1224-1231, 1974.

Edwards, J, H. Wagner, and W. Wood, "Blue Bell trims its inventory," *Interfaces*, Vol. 15, No. 1, pp. 34-52, 1985.

Lewis, T., "Personal Operations Research: Practicing OR on Ourselves," *Interfaces*, Vol. 26, No. 5, pp. 34-41, 1996.

Nahmias, S., *Production and Operations Analysis*, 5th ed., Irwin, Homewood, IL, 2005.

Silver, E., D. Pyke, and R. Peterson, *Decision Systems for Inventory Management and Production Control*, 3rd ed., Wiley, New York, 1998.

Tersine, R., *Principles of Inventory and Materials Management*, 3rd ed., North Holland, New York, 1988.

Waters, C., *Inventory Control and Management*, Wiley, New York, 1992.

第11章　决策分析与对策

本章导读　在实际中常常会出现涉及有限多个可能方案的决策问题, 用来解决这些问题的工具和已有数据的类型 (如确定型、概率型或不确定型) 有着很大的关系. 层次分析法 (Analytic Hierarchy Process, AHP) 是处理确定性决策问题的一种最常用的工具, 它将主观判断按照合理的方式进行量化, 用作最终决策的依据. 对概率型数据, 用来比较各种不同可能方案期望费用 (或利润) 的决策树成为决策的基础, 而不确定因素下的决策则是利用多个指标来反映决策者对于风险从乐观到悲观的态度. 对策论是不确定因素下决策的另一种工具, 在对策论中, 具有冲突目标的两个对手都试图从对各自最不利的条件下获得最好的结果. 为了说明这些工具对于解决实际问题的重要性, 第 24 章给出了 4 个案例, 分别是用 AHP 方法解决 CIM 实验室的布局规划问题, 用决策树分析确定旅店订房的上限控制, 用贝叶斯概率评估医疗化验结果, 以及利用对策论对参加莱德杯 (Ryder Cup) 比赛的高尔夫球手进行打分. 为了帮助大家详细了解这些不同的工具, 本章提供了 4 个电子表格程序. 你还会发现, TORA 程序对于对策模型的图形和代数求解非常有用. 本章的学习需要用到概率统计的基本知识.

本章共包括 4 个实际应用的简介、10 个例题、4 个 Excel 程序、63 个节后习题以及 5 个案例. 这些案例收录在附录 E 中. AMPL/Excel/Solver/TORA 程序都存在下载资料 ch11Files 文件夹下.

实际应用问题——计算机集成制造(CIM)设施的布局规划

某学术机构的工程学院想要在一幢腾空的大楼内建一个 CIM 实验室. 这个新的实验室将作为教学和科研设施, 并作为一个产业技术进步中心. 学院征集了有关新建实验室布局规划的建议, 从这些建议方案中汇集了每个建设单元理想情况下以及最低要求情况下的面积数, 运用了 AHP 和目标规划方法, 得到了能够满足教学、科研和产业服务需要的一个令人满意的折衷解决方案. 这项研究的细节请见第 24 章的案例 9.

11.1　确定型决策——层次分析法 (AHP)

第 2 章至第 8 章介绍的线性规划模型都是确定型决策的例子, 其中所有的函数都有确切的定义. 在运用 AHP 时, 我们要对影响决策过程的各种想法、感觉以

及情绪进行量化, 以提供一种数量化尺度, 评判各种行动方案的重要性.

例 11.1-1 (**AHP的总体思路**)

Martin Hans 是个聪明的高中毕业生, 他从 A, B, C 这 3 所大学申请到了全额奖学金. 为了最终选择一所大学, Martin 指定了两个主要指标准则: 地点和学术名气. 作为一名优秀生, 他认为大学的学术名气要比地点重要 5 倍, 所以对地点的权重给了大约 17%, 而给名气的权重为 83%. 然后他采用了一种系统化的分析方法 (下面要介绍) 对这 3 所大学从地点和学术名气的角度进行排序. 下表给出了这 3 所大学的两个指标的结果:

指标准则	百分比权重的估计		
	A 大学	B 大学	C 大学
地点	12.9	27.7	59.4
名气	54.5	27.3	18.2

图 11.1 给出了这个决策问题的结构, 包括了 1 个层次 (办学水平)、2 个准则 (地点和名气) 和 3 个可能的决策方案 (大学 A, B, C).

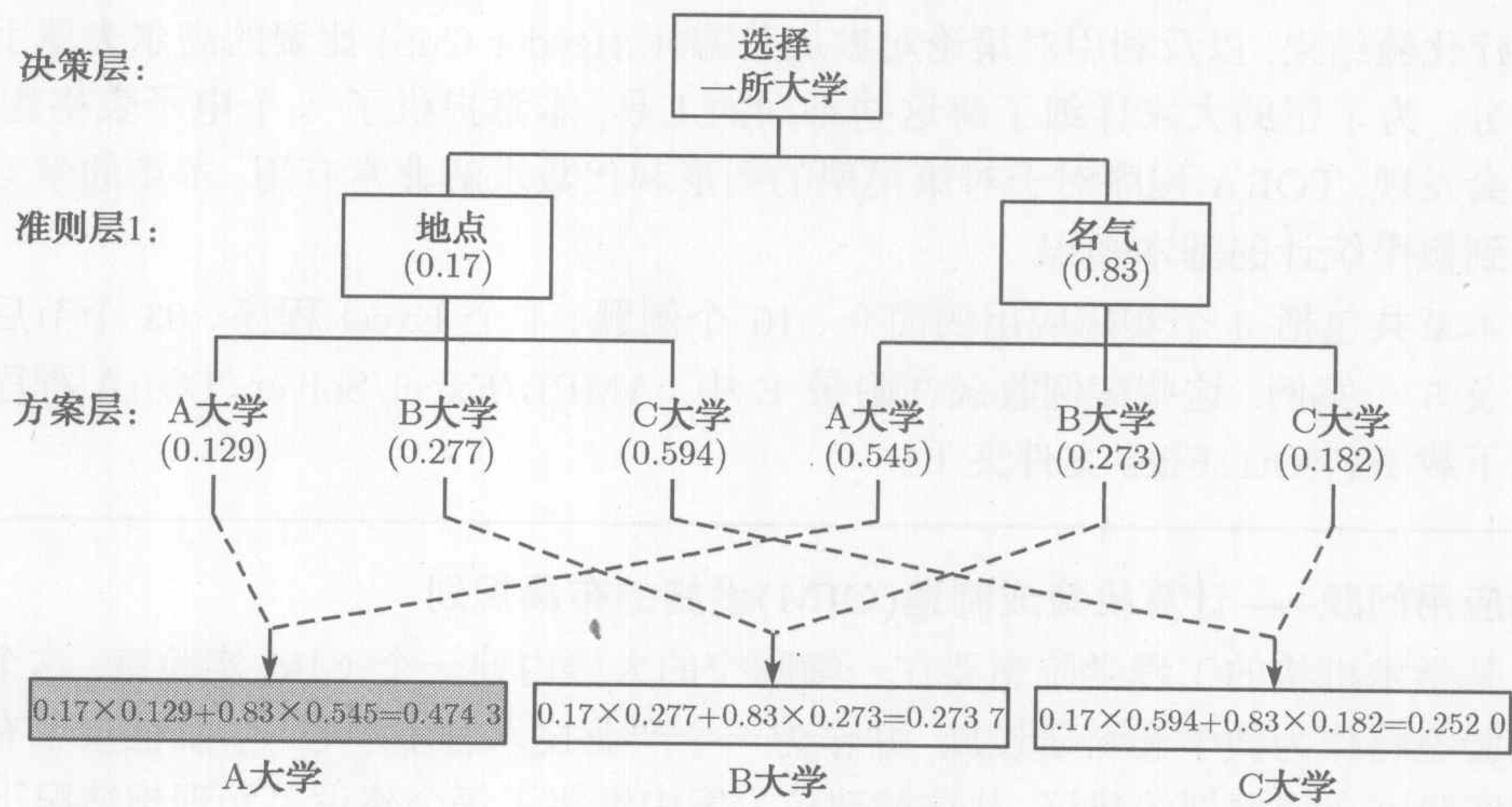


图 11.1　例 11.1-1 的 AHP 计算结果

根据下面复合权重的计算结果, 对每所大学进行评定.

$$\text{A大学} = 0.17 \times 0.129 + 0.83 \times 0.545 = \mathbf{0.474\,3}$$

$$\text{B大学} = 0.17 \times 0.277 + 0.83 \times 0.273 = 0.273\,7$$

$$\text{C大学} = 0.17 \times 0.594 + 0.83 \times 0.182 = 0.252\,0$$

根据上述计算结果, A 大学的复合权重最高, 因此代表了 Martin 的最佳选择.

评注 AHP 的一般结构可以包括多个层次的准则. 假设在例 11.1-1 中, Martin 的同胞胎妹妹 Jane 同样收到了这 3 所大学的奖学金, 他们的父母规定, 为了共用一部汽车他们俩必须上同一所大学. 图 11.2 显示了这一决策问题, 它的指标有两层. 第 1 层的 p 和 q 的值 (假设相等) 代表 Martin 和 Jane 的行动方案关于选择过程的相对权重, 第 2 层用权重 (p_1, p_2) 和 (q_1, q_2) 来反映 Martin 和 Jane 对每所大学地点和名气的偏好. 可以类似地来解释这个决策图上的其他部分. 注意到 $p+q=1, p_1+p_2=1, q_1+q_2=1, p_{11}+p_{12}+p_{13}=1, p_{21}+p_{22}+p_{23}=1, q_{11}+q_{12}+q_{13}=1, q_{21}+q_{22}+q_{23}=1$. 图 11.2 中显示了如何进行 A 大学复合权重的计算.

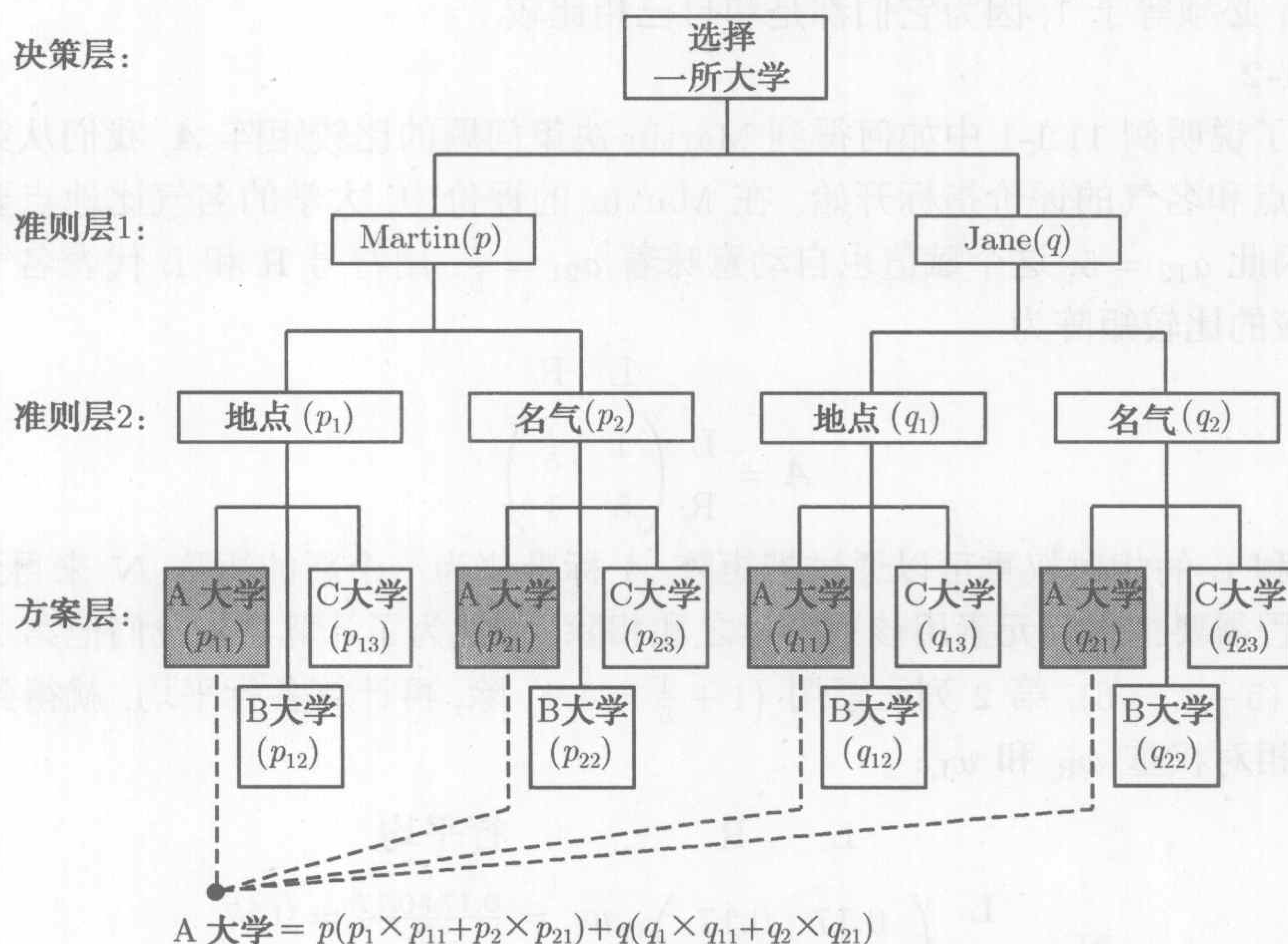


图 11.2 对例 11.1-1 决策问题的修正

习题 11.1A

*1. 假设对 Martin 和 Jane 的问题给出下列权重:

$$p=0.5,\ q=0.5$$
$$p_1=0.17,\ p_2=0.83$$
$$p_{11}=0.129,\ p_{12}=0.277,\ p_{13}=0.594$$
$$p_{21}=0.545,\ p_{22}=0.273,\ p_{23}=0.182$$
$$q_1=0.3,\ q_2=0.7$$
$$q_{11}=0.2,\ q_{12}=0.3,\ q_{13}=0.5$$
$$q_{21}=0.5,\ q_{22}=0.2,\ q_{23}=0.3$$

根据这些信息, 对这 3 所大学进行排序.

确定权重 AHP 的关键是确定相对权重 (如例 11.1-1 中用到的权值), 才能对可能的决策行动方案进行排序. 假定我们正在处理某个给定层次的 n 个评价指标, 我们需要建立一个 $n \times n$ 的成对**比较矩阵**(comparison matrix) $\boldsymbol{A}$, 对决策者给出的不同指标的相对重要性评价加以量化. 进行成对比较是要让第 i 行 $(i = 1, 2, \cdots, n)$ 的指标相对于每个其他指标进行度量. 令 a_{ij} 为 $\boldsymbol{A}$ 的第 (i, j) 元素, AHP 采用一种从 1~9 的离散尺度进行度量, $a_{ij} = 1$ 代表 i 和 j 是同等重要的, $a_{ij} = 5$ 说明 i 比 j 要重要得多, $a_{ij} = 9$ 代表 i 和 j 相比绝对重要, 1~9 的其他中间值可相应解释. 评价中的一致性要求是, 如果 $a_{ij} = k$ 则自动意味着 $a_{ji} = \frac{1}{k}$; 还有, $\boldsymbol{A}$的所有对角线元素 a_{ii} 必须等于 1, 因为它们都是和自己相比较.

例 11.1-2

为了说明例 11.1-1 中如何得到 Martin 决策问题的比较矩阵 $\boldsymbol{A}$, 我们从第一层大学地点和名气的评价指标开始. 在 Martin 的评价中, 大学的名气比地点要重要得多, 因此 $a_{12} = 5$, 这个赋值也自动意味着 $a_{21} = \frac{1}{5}$. 用符号 R 和 L 代表名气和地点, 相应的比较矩阵为

$$\boldsymbol{A} = \begin{matrix} & \begin{matrix} \mathrm{L} & \mathrm{R} \end{matrix} \\ \begin{matrix} \mathrm{L} \\ \mathrm{R} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{5} \\ 5 & 1 \end{pmatrix} \end{matrix}$$

R 和 L 的相对权重可以通过把矩阵 $\boldsymbol{A}$ 标准化为一个新的矩阵 $\boldsymbol{N}$ 来得到. 标准化过程需要把每列元素用该列元素之和相除. 因此为了计算 $\boldsymbol{N}$, 我们把第 1 列元素除以 $(5 + 1 = 6)$, 第 2 列元素用 $(1 + \frac{1}{5} = 1.2)$ 除, 再计算各行平均, 就得到了所需要的相对权重 w_{R} 和 w_{L}:

$$\boldsymbol{N} = \begin{matrix} & \begin{matrix} \mathrm{L} & \mathrm{R} \end{matrix} & \text{行平均} \\ \begin{matrix} \mathrm{L} \\ \mathrm{R} \end{matrix} & \begin{pmatrix} 0.17 & 0.17 \\ 0.83 & 0.83 \end{pmatrix} & \begin{matrix} w_{\mathrm{L}} = \frac{0.17+0.17}{2} = 0.17 \\ w_{\mathrm{R}} = \frac{0.83+0.83}{2} = 0.83 \end{matrix} \end{matrix}$$

上述计算得到图 11.1 中所用的 $w_{\mathrm{L}} = 0.17, w_{\mathrm{R}} = 0.83$. $\boldsymbol{N}$的两个列是相同的, 只有当决策者的比较矩阵$\boldsymbol{A}$达到了完全一致的时候才有这个性质. 本节的后面还要进一步讨论这一点.

对选择大学 A, B, C 的每个评价指标 L 和 R 的相对权重, 可通过下面两个比较矩阵计算, 其元素根据 Martin 对这 3 所大学的相对重要性判断得出.

$$\boldsymbol{A}_{\mathrm{L}} = \begin{matrix} & \begin{matrix} \mathrm{A} & \mathrm{B} & \mathrm{C} \end{matrix} \\ \begin{matrix} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{5} \\ 2 & 1 & \frac{1}{2} \\ 5 & 2 & 1 \end{pmatrix} \end{matrix}, \quad \boldsymbol{A}_{\mathrm{R}} = \begin{matrix} & \begin{matrix} \mathrm{A} & \mathrm{B} & \mathrm{C} \end{matrix} \\ \begin{matrix} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{matrix} & \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{2} & 1 & \frac{3}{2} \\ \frac{1}{3} & \frac{2}{3} & 1 \end{pmatrix} \end{matrix}$$

对各列求和可以得到

$$A_L\text{列求和} = (8, 3.5, 1.7)$$
$$A_R\text{列求和} = (1.83, 3.67, 5.5)$$

分别用列求和结果除所有元素得到以下标准化的矩阵:

$$N_L = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0.125 & 0.143 & 0.118 \\ 0.250 & 0.286 & 0.294 \\ 0.625 & 0.571 & 0.588 \end{pmatrix} \end{array} \quad \begin{array}{l} \text{行平均} \\ w_{LA} = \frac{0.125+0.143+0.118}{3} = 0.129 \\ w_{LB} = \frac{0.250+0.286+0.294}{3} = 0.277 \\ w_{LC} = \frac{0.625+0.571+0.588}{3} = 0.594 \end{array}$$

$$N_R = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0.545 & 0.545 & 0.545 \\ 0.273 & 0.273 & 0.273 \\ 0.182 & 0.182 & 0.182 \end{pmatrix} \end{array} \quad \begin{array}{l} \text{行平均} \\ w_{RA} = \frac{0.545+0.545+0.545}{3} = 0.545 \\ w_{RB} = \frac{0.273+0.273+0.273}{3} = 0.273 \\ w_{RC} = \frac{0.182+0.182+0.182}{3} = 0.182 \end{array}$$

向量值 $(w_{LA}, w_{LB}, w_{LC}) = (0.129, 0.277, 0.594)$ 提供了大学 A, B, C 分别的地点权重. 类似地, $(w_{RA}, w_{RB}, w_{RC}) = (0.545, 0.273, 0.182)$ 给出了关于学术名气的相对权重.

比较矩阵的一致性 在例 11.1-2 中, 经过标准化的矩阵 $\boldsymbol{N}$ 和 $\boldsymbol{N}_R$ 的所有列都一样的, 但 $\boldsymbol{N}_L$ 的各列不一样. 在这种情况下, 我们称原来的比较矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}_R$ 是一致的, 但 $\boldsymbol{A}_L$ 是不一致的.

一致性意味着决策者关于成对比较的判断是一致的. 数学上, 我们说比较矩阵 $\boldsymbol{A}$ 是一致的, 如果

$$a_{ij}a_{jk} = a_{ik}, \quad \text{对所有的} i, j, k$$

例如在例 11.1-2 中, $a_{13} = 3$ 并且 $a_{12}a_{23} = 2 \times \frac{3}{2} = 3$. 这一性质要求, $\boldsymbol{A}$ 的所有列 (和行) 都是线性相关的. 特别地, 任何 2×2 比较矩阵的各列根据定义是相关的, 因此 2×2 矩阵的列总是一致的.

然而, 通常情况下比较矩阵并不都是一致的, 因为构造这些矩阵是根据人为的判断, 我们接受和容许某种"适当"程度的不一致性存在.

为了确定某一水平的不一致性是否"适当", 需要建立比较矩阵 $\boldsymbol{A}$ 的一个定量性的度量. 我们在例 11.1-2 中已经看到, 一个完全一致的 $\boldsymbol{A}$ 会产生一个标准化的矩阵 $\boldsymbol{N}$, 其中所有的列是完全相同的, 即

$$\boldsymbol{N} = \begin{pmatrix} w_1 & w_1 & \cdots & w_1 \\ w_2 & w_2 & \cdots & w_2 \\ \vdots & \vdots & & \vdots \\ w_n & w_n & \cdots & w_n \end{pmatrix}$$

由此可知, 原来的比较矩阵 $\boldsymbol{A}$ 可以从 $\boldsymbol{N}$ 通过用 w_i 除以第 i 列的各个元素得到. 因此有

$$\boldsymbol{A}=\begin{pmatrix} 1 & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & 1 \end{pmatrix}$$

由上述 $\boldsymbol{A}$ 的定义得

$$\begin{pmatrix} 1 & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & 1 \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}=\begin{pmatrix} nw_1 \\ nw_2 \\ \vdots \\ nw_n \end{pmatrix}=n\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

简洁地表示为, 设 $\boldsymbol{w}$ 表示相对权重 $w_i, i=1,2,\cdots,n$ 的列向量, $\boldsymbol{A}$ 是一致的, 如果有

$$\boldsymbol{A}\boldsymbol{w}=n\boldsymbol{w}$$

对于 $\boldsymbol{A}$ 不一致的情况, 相对权重 w_i 用标准化矩阵 $\boldsymbol{N}$ 第 i 行 n 个元素的平均值近似 (见例 11.1-2). 令 $\overline{\boldsymbol{w}}$ 为算出来的平均向量, 可以证明

$$\boldsymbol{A}\overline{\boldsymbol{w}}=n_{\max}\overline{\boldsymbol{w}}, \quad n_{\max}\geqslant n$$

在这种情况下, $n_{\max}$ 越接近于 n, 比较矩阵 $\boldsymbol{A}$ 就越一致. 根据这一现象, AHP 计算下面的**一致性比**(consistency ratio):

$$CR=\frac{CI}{RI}$$

其中

$$CI=\boldsymbol{A}\text{的一致性指数}=\frac{n_{\max}-n}{n-1}$$

$$RI=\boldsymbol{A}\text{的随机一致性}=\frac{1.98(n-2)}{n}$$

随机一致性指数 RI 的计算, 根据经验为随机产成的比较矩阵 $\boldsymbol{A}$ 的一个大样本平均 CI.

如果 $CR\leqslant 0.1$, 则一致性程度是可以接受的, 否则认为不一致性太高, 决策者需要重新估计 $\boldsymbol{A}$ 的元素 a_{ij}, 以便得到更好的一致性.

我们从 $\boldsymbol{A}\overline{\boldsymbol{w}}=n_{\max}\overline{\boldsymbol{w}}$ 计算 $n_{\max}$ 的值, 并注意到其第 i 个方程是

$$\sum_{j=1}^{n} a_{ij}\overline{w}_j=n_{\max}\overline{w}_i, \quad i=1,2,\cdots,n$$

由 $\sum_{i=1}^{n}\overline{w_i}=1$, 我们得到

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}\overline{w}_j\right)=n_{\max}\sum_{i=1}^{n}\overline{w}_i=n_{\max}$$

这意味着, $n_{\max}$ 的值可以通过先计算列向量 $\boldsymbol{A}\overline{\boldsymbol{w}}$, 然后对它的各元素求和来得到.

例 11.1-3

在例 11.1-2 中, 矩阵 $\boldsymbol{A}_{\mathrm{L}}$ 不一致, 因为它的 $\boldsymbol{N}_{\mathrm{L}}$ 各列不相同. 检验 $\boldsymbol{N}_{\mathrm{L}}$ 的不一致性程度.

首先计算 $n_{\max}$. 从例 11.1-2 可得

$$\overline{w}_1=0.129,\quad \overline{w}_2=0.277,\quad \overline{w}_3=0.594$$

因此,

$$\boldsymbol{A}_{\mathrm{L}}\overline{\boldsymbol{w}}=\begin{pmatrix}1 & \frac{1}{2} & \frac{1}{5}\\ 2 & 1 & \frac{1}{2}\\ 5 & 2 & 1\end{pmatrix}\begin{pmatrix}0.129\\0.277\\0.594\end{pmatrix}=\begin{pmatrix}0.386\ 3\\0.832\ 0\\1.793\ 0\end{pmatrix}$$

由此得出

$$n_{\max}=0.386\ 3+0.832\ 0+1.793\ 0=3.011\ 3$$

因此, 对于 $n=3$,

$$CI=\frac{n_{\max}-n}{n-1}=\frac{3.011\ 3-3}{3-1}=0.005\ 65$$

$$RI=\frac{1.98(n-2)}{n}=\frac{1.98\times 1}{3}=0.66$$

$$CR=\frac{CI}{RI}=\frac{0.005\ 65}{0.66}=0.008\ 56$$

由于 $CR<0.1$, $\boldsymbol{A}_{\mathrm{L}}$ 中的不一致性程度是可以接受的.

Excel 程序

程序 ExcelAHP.xls 用来处理最大 8×8 的比较矩阵. 与第 9 章和第 10 章中的模型一样, 用户输入后启动模型计算. 图 11.3 显示了该程序对例 11.1-2① 模型的计算. 这个问题的比较矩阵在电子表格的数据输入区 (上方) 一次输入一个, 虽然分层输入符合我们的自然习惯, 然而比较矩阵的输入顺序并不重要. 在对一个比较矩阵输入数据时, 电子表格的输出区域 (下方) 会提供相应的标准化矩阵以及相应的一

① 由于人工舍入误差的计算, 电子表格更加精确的结果与例 11.1-2 和例 11.1-3 的结果有所差别. 为了节省空间, 列 F:I 和行 11:13 被压缩了.

致性比系数 CR. 用户必须从列 J 复制权重 w 并粘贴到解汇总区 (电子表格的右方). 记住在执行这一步操作时, 用菜单命令 Paste Special ⇒ Values, 记住保存永久性纪录. 重复这一过程, 直至把所有的比较矩阵都储存在 K:R 列.

	A	B	C	D	E	J	K	L	M	N
1	AHP-Analytic Hierarchy Process									
2	Input: Comparison matrix						Solution summary			
3	Matrix name:	AL						A		
4	Matrix size=	3	<<Maximum 8				R	0.83333		
5	Matrix data:	UA	UB	UC			L	0.16667		
6	UA	1	0.5	0.2				AR		AL
7	UB	2	1	0.5			UA	0.54545	UA	0.1285
8	UC	5	2	1			UB	0.27273	UB	0.27661
9							UC	0.18182	UC	0.59489
13										
14	Col sum	8	3.5	1.7						
15	Output: Normalized martix									
16		nMax=	3.00746	CR=	0.0056					
17		UA	UB	UC		Weight				
18	UA	0.12500	0.14286	0.11765		0.12850				
19	UB	0.25000	0.28571	0.29412		0.27661	Final ranking			
20	UC	0.62500	0.57143	0.58824		0.59489		UA= 0.47596		
21								UB= 0.27337		
22								UC= 0.25066		
26										

图 11.3 例 11.1-2 的 Excel 求解 (文件 excelAHP. xls)

图 11.3 中的单元格 (K20:K22) 给出了最终的打分结果. 单元格 K20 的公式为

$$= \$L\$4 * \$L7 + \$L\$5 * \$N7$$

这个公式提供了对备选方案 A 大学的最终评估结果, 并复制到单元格 K21 和 K22, 对大学 B 和大学 C 进行评估. 请注意 K20 中的公式是如何构造的: 对方案 A 大学的单元格引用必须是列固定的 (即, $L7 和 $N7), 而其他所有的引用必须是行和列固定的 (即, $L$4 和 $L$5). 所复制的公式的正确性检验要求, 每个矩阵 (列固定) 的*备选方案*的权重都出现在同一列, 中间没有空白单元. 例如, 在图 11.3 中, 列 L 中的 AR 权重不能在两列之间间断. 对 N 列中的 AL 权重亦是如此. 对 A 权重的位置没有限制, 因为它们在公式中是行和列固定的.

你可以对该公式进行修正, 来直接反映备选方案的名称. 例如 A 大学备选方案的公式输入为 = $K7&“=”&TEXT($ L$4*$L7+$L$5*$N7,“####0.000 00”).

上述用来评估各备选方案的程序能够扩展到任意层次数. 一旦你对第一个备选方案建立了正确的公式, 就可以把同样的公式用到剩下的备选方案上去, 只需要把它复制到 (同一列) 下边的行上. 记住, 公式中除了对可选方案的引用必须是列固定的以外, *所有的*其他引用必须是行和列都固定的. 习题 11.1B 中的第 1 题要求你建立一个 3 层次问题的公式.

习题 11.1B[①]

1. 考虑习题 11.1A 第 1 题的数据. 按照合理的顺序把权重复制到电子表格程序 excelAHP.xls 的解汇总区域, 然后建立评估第一个可选方案 A 大学的公式, 并复制这个公式来评估余下的两个可选方案.

***2.** C&H 的人事部门已将招聘一名员工的范围缩小到 3 位候选人：Steve (S)、Jane (J) 和 Maisa (M). 最终的选择要根据 3 项评价指标：面试结果 (I), 工作经历 (E) 以及推荐意见 (R). 该部门利用一个矩阵 $\boldsymbol{A}$(下面给出) 在这 3 项指标中评出倾向性意见. 在对 3 位候选人面试并整理出有关他们的工作经历和推荐意见以后, 建立了 3 个矩阵 $\boldsymbol{A}_{\rm J}$、$\boldsymbol{A}_{\rm E}$ 和 $\boldsymbol{A}_{\rm R}$. 应该聘用哪一位候选人呢? 并评价数据的一致性.

$$\boldsymbol{A}=\begin{array}{c} \\ \text{I} \\ \text{E} \\ \text{R} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{I} & \text{E} & \text{R}\end{array}\\ \begin{pmatrix} 1 & 2 & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{5} \\ 4 & 5 & 1 \end{pmatrix}\end{array} \qquad \boldsymbol{A}_{\rm I}=\begin{array}{c} \\ \text{S} \\ \text{J} \\ \text{M} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{S} & \text{J} & \text{M}\end{array}\\ \begin{pmatrix} 1 & 3 & 4 \\ \frac{1}{3} & 1 & \frac{1}{5} \\ \frac{1}{4} & 5 & 1 \end{pmatrix}\end{array}$$

$$\boldsymbol{A}_{\rm E}=\begin{array}{c} \\ \text{S} \\ \text{J} \\ \text{M} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{S} & \text{J} & \text{M}\end{array}\\ \begin{pmatrix} 1 & \frac{1}{3} & 2 \\ 3 & 1 & \frac{1}{2} \\ \frac{1}{2} & 2 & 1 \end{pmatrix}\end{array} \qquad \boldsymbol{A}_{\rm R}=\begin{array}{c} \\ \text{S} \\ \text{J} \\ \text{M} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{S} & \text{J} & \text{M}\end{array}\\ \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 2 & 1 & \frac{1}{2} \\ 1 & 2 & 1 \end{pmatrix}\end{array}$$

3. Kevin 和 June Park (K&J) 夫妇正在考虑买一幢新房子, 有 A, B, C 这 3 幢房子可选. Park 夫妇同意用两个指标来选房：庭院布置 (Y) 和离工作单位的距离 (W). 并建立了以下比较矩阵. 请按照优先级对这 3 幢房子进行排队, 并计算每个矩阵的一致性比.

$$\boldsymbol{A}=\begin{array}{c} \\ \text{K} \\ \text{J} \end{array}\!\!\begin{array}{c}\begin{array}{cc}\text{K} & \text{J}\end{array}\\ \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & 1 \end{pmatrix}\end{array}$$

$$\boldsymbol{A}_{\rm K}=\begin{array}{c} \\ \text{Y} \\ \text{W} \end{array}\!\!\begin{array}{c}\begin{array}{cc}\text{Y} & \text{W}\end{array}\\ \begin{pmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{pmatrix}\end{array} \qquad \boldsymbol{A}_{\rm J}=\begin{array}{c} \\ \text{Y} \\ \text{W} \end{array}\!\!\begin{array}{c}\begin{array}{cc}\text{Y} & \text{W}\end{array}\\ \begin{pmatrix} 1 & 4 \\ \frac{1}{4} & 1 \end{pmatrix}\end{array}$$

$$\boldsymbol{A}_{\rm KY}=\begin{array}{c} \\ \text{A} \\ \text{B} \\ \text{C} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{A} & \text{B} & \text{C}\end{array}\\ \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{2} & 1 & 2 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{pmatrix}\end{array} \qquad \boldsymbol{A}_{\rm KW}=\begin{array}{c} \\ \text{A} \\ \text{B} \\ \text{C} \end{array}\!\!\begin{array}{c}\begin{array}{ccc}\text{A} & \text{B} & \text{C}\end{array}\\ \begin{pmatrix} 1 & 2 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{3} \\ 2 & 3 & 1 \end{pmatrix}\end{array}$$

① 应该用电子表格程序 excelAHP.xls 来检验你的计算.

$$A_{JY}=\begin{matrix} & A & B & C\\ A & 1 & 4 & 2\\ B & \frac{1}{4} & 1 & 3\\ C & \frac{1}{2} & \frac{1}{3} & 1\end{matrix} \qquad A_{JW}=\begin{matrix} & A & B & C\\ A & 1 & \frac{1}{2} & 4\\ B & \frac{1}{2} & 1 & 3\\ C & \frac{1}{4} & \frac{1}{3} & 1\end{matrix}$$

***4.** 一位新作者设定了 3 个指标来选择一家出版社出版一本运筹学教材：版税比例 (R)、市场销售 (M) 和提前支付 (A). 有两家出版商 H 和 P 对这本书感兴趣. 请利用下列比较矩阵对这两家出版商进行排队，并评价决策的一致性程度.

$$A=\begin{matrix} & R & M & A\\ R & 1 & 1 & \frac{1}{4}\\ M & 1 & 1 & \frac{1}{5}\\ A & 4 & 5 & 1\end{matrix}$$

$$A_R=\begin{matrix} & H & P\\ H & 1 & 2\\ P & \frac{1}{2} & 1\end{matrix} \qquad A_M=\begin{matrix} & H & P\\ H & 1 & \frac{1}{2}\\ P & 2 & 1\end{matrix} \qquad A_A=\begin{matrix} & H & P\\ H & 1 & 1\\ P & 1 & 1\end{matrix}$$

5. 某政治学教授想要预测学校董事会的选举结果. 有 3 个候选人 Ivy (I), Bahrn (B) 和 Smith (S) 竞争一个席位. 这位教授把投票人划分成 3 类：左 (L)、中 (C)、右 (R). 用 3 个因素对候选人做出评价：受教育经历 (E). 对某些问题的立场 (S) 和个人性格 (P). 以下是左、中、右的第一层比较矩阵.

$$A=\begin{matrix} & L & C & R\\ L & 1 & 2 & \frac{1}{2}\\ C & \frac{1}{2} & 1 & \frac{1}{5}\\ R & 2 & 5 & 1\end{matrix} \qquad A_L=\begin{matrix} & E & S & P\\ E & 1 & 3 & \frac{1}{2}\\ S & \frac{1}{3} & 1 & \frac{1}{3}\\ P & 2 & 3 & 1\end{matrix}$$

$$A_C=\begin{matrix} & E & S & P\\ E & 1 & 2 & 2\\ S & \frac{1}{2} & 1 & 1\\ P & \frac{1}{2} & 1 & 1\end{matrix} \qquad A_R=\begin{matrix} & E & S & P\\ E & 1 & 1 & 9\\ S & 1 & 1 & 8\\ P & \frac{1}{9} & \frac{1}{8} & 1\end{matrix}$$

该教授还生成了第 2 层的 9 个比较矩阵代表 3 个候选人的经历、立场和个人性格，然后用 AHP 方法把这些矩阵简化成了下面的相对权重表：

候选人	左			中			右		
	E	S	P	E	S	P	E	S	P
Ivy	0.1	0.2	0.3	0.3	0.5	0.2	0.7	0.1	0.3
Bahrn	0.5	0.4	0.2	0.4	0.2	0.4	0.1	0.4	0.2
Smith	0.4	0.4	0.5	0.3	0.3	0.4	0.2	0.5	0.5

请确定胜出的候选人, 并评价决策的一致性.

6. 某学校园区为了满足新的预算限制, 急需削减小学的支出. 有两种备选方案: 取消体育课 (E), 或者取消音乐课 (M). 园区主管成立了一个由学校董事会 (S) 和家长协会 (P) 等额代表组成的委员会, 来研究这一情况并提出建议. 该委员会决定从预算限制 (B) 和学生需要 (N) 两个方面来研究这一问题. 经过分析产生了下列比较矩阵:

$$\boldsymbol{A}_{\mathrm{S}} = \begin{matrix} & \begin{matrix} \mathrm{B} & \mathrm{N} \end{matrix} \\ \begin{matrix} \mathrm{B} \\ \mathrm{N} \end{matrix} & \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \end{matrix} \quad \boldsymbol{A}_{\mathrm{P}} = \begin{matrix} & \begin{matrix} \mathrm{B} & \mathrm{N} \end{matrix} \\ \begin{matrix} \mathrm{B} \\ \mathrm{N} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 1 \end{pmatrix} \end{matrix}$$

$$\boldsymbol{A}_{\mathrm{SB}} = \begin{matrix} & \begin{matrix} \mathrm{E} & \mathrm{M} \end{matrix} \\ \begin{matrix} \mathrm{E} \\ \mathrm{M} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 1 \end{pmatrix} \end{matrix} \quad \boldsymbol{A}_{\mathrm{SN}} = \begin{matrix} & \begin{matrix} \mathrm{E} & \mathrm{M} \end{matrix} \\ \begin{matrix} \mathrm{E} \\ \mathrm{M} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{pmatrix} \end{matrix}$$

$$\boldsymbol{A}_{\mathrm{PB}} = \begin{matrix} & \begin{matrix} \mathrm{E} & \mathrm{M} \end{matrix} \\ \begin{matrix} \mathrm{E} \\ \mathrm{M} \end{matrix} & \begin{pmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{pmatrix} \end{matrix} \quad \boldsymbol{A}_{\mathrm{PN}} = \begin{matrix} & \begin{matrix} \mathrm{E} & \mathrm{M} \end{matrix} \\ \begin{matrix} \mathrm{E} \\ \mathrm{M} \end{matrix} & \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & 1 \end{pmatrix} \end{matrix}$$

请分析这个决策问题并提出建议.

7. 某人正在考虑购买一辆汽车, 已经把选择范围局限在 3 款车型之内: M1, M2, M3. 购车的决定因素包括: 购买价格 (PP)、维修费用 (MC)、城内驾驶成本 (CD) 以及乡村驾驶成本 (RD). 下表提供了 3 种车型的相关数据:

车型	PP($)	MC($)	CD($)	RD($)
M1	6 000	1 800	4 500	1 500
M2	8 000	1 200	2 250	750
M3	10 000	600	1 125	600

利用这些费用数据建立比较矩阵. 评价这些矩阵的一致性, 并确定所要购买的车型.

11.2 风险型决策

在风险条件下, 我们用概率分布来描述相对每个行动方案的益损. 因此, 风险型决策可以根据**期望值评估指标**来作出, 利用这一指标对各个决策方案根据最大期望收益或最小期望费用进行比较. 但是这一方法也有局限性, 可以对期望值指标进行修改, 使其适应于其他情况.

实际应用——旅馆订房上限

La Posada 旅店共有 300 间房间. 旅店的客人既有出差人士, 也有度假旅游者. 这些房间可以事先按折扣价预订 (通常订给旅游者), 商务出差旅客的订房时间通常比较晚, 他们支付全价. 因此 La Posada 旅店必须为休闲旅客的折扣订房数量设

定一个*定房上限*, 以便预留一定数量的客房给付全价的商务旅客. 在第 24 章的案例 10 中, 利用了决策树分析方法来确定这个订房上限.

11.2.1　基于决策树的期望值指标

期望值指标寻求期望 (平均) 利润的最大化, 或者期望费用的最小化. 这类问题的数据假定, 相应每个决策方案的益损 (或费用) 都是随机性的.

决策树分析　下面的例子考虑带有有限个决策方案以及明确收益矩阵的简单决策问题.

例 11.2-1

假定你想要在股票市场投资 \$10 000 购买 A, B 两家公司其中一家的股份. 公司 A 的股票虽然有风险, 但可在下一年获得 50% 的投资回报. 假如股票市场条件不利的话 (即熊市), 这只股票可能损失价值的 20%. 公司 B 提供一种安全的投资, 牛市的投资回报率为 15%, 而熊市只有 5% 的回报. 你所看到的所有出版物 (年底总是有大量的这类出版物!) 都预测, 牛市的机会是 60%, 而熊市的机会是 40%. 在这种情况下, 你应该把资金投向哪里呢?

这一决策问题可以总结如下:

可能的决策方案	\$10 000 投资的一年回报	
	牛市 (\$)	熊市 (\$)
公司 A 的股票	5 000	−2 000
公司 B 的股票	1 500	500
发生的概率	0.6	0.4

这个问题也可以表示为一个**决策树** (decision tree), 如图 11.4 所示. 树中用了两种节点: 正方形 (□) 代表一个*决策点*, 圆圈 (○) 代表一个*机会事件*. 这样, 从决策点 1 分出两个分支, 代表投资股票 A 和股票 B 的两种可能方案. 接下来又从机会事件 2 和 3 分别分出两个分叉, 代表牛市和熊市, 并标出了各自的概率和益损值.

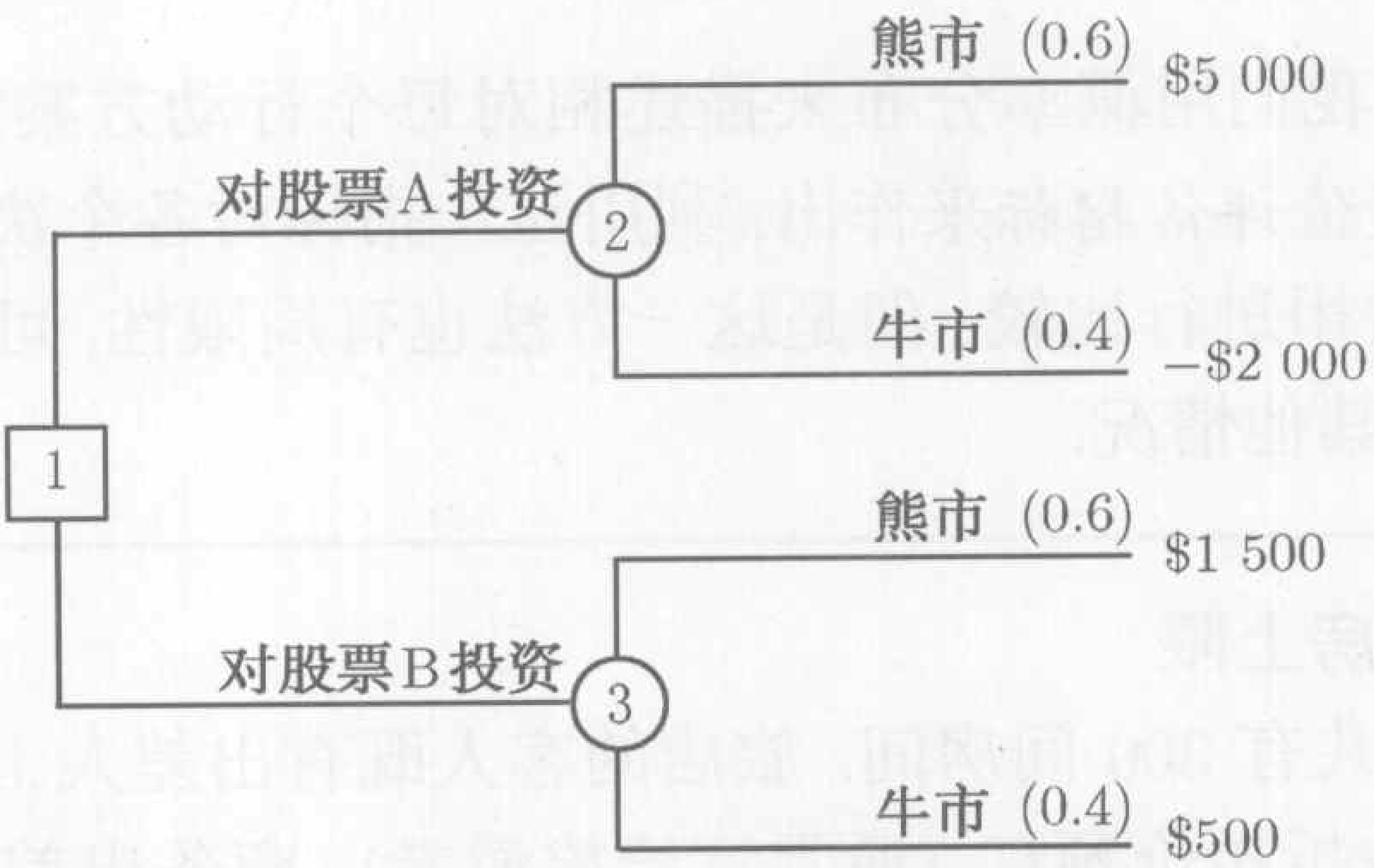


图 11.4　股票市场问题的决策树表示

从图 11.4 中, 这两种可能方案的 1 年期望回报分别为

$$股票A = \$5\ 000 \times 0.6 + (-2\ 000) \times 0.4 = \$2\ 200$$
$$股票B = \$1\ 500 \times 0.6 + \$500 \times 0.4 = \$1\ 100$$

根据上述计算结果, 你的决策是在股票 A 上投资.

评注 按照决策理论的术语, 前面例子中的牛市和熊市是指**自然状态** (state of nature), 它们发生的机会是随机 (概率性) 的 (0.6 对 0.4). 一般来说, 一个决策问题可能包括 n 个自然状态和 m 个可选方案. 假定 $p_j(>0)$ 是自然状态 j 发生的概率, a_{ij} 为给定的自然状态 j 时方案 i 的益损值 ($i=1,2,\cdots,m; j=1,2,\cdots,n$), 则可能方案 i 的期望益损值可计算为

$$\mathrm{EV}_i = a_{i1}p_1 + a_{i2}p_2 + \cdots + a_{in}p_n, \quad i = 1, 2, \cdots, n$$

按照定义, $p_1 + p_2 + \cdots + p_n = 1$.

最好的方案是, 其相应的 $\mathrm{EV}^* = \max_i\{\mathrm{EV}_i\}$, 或者 $\mathrm{EV}^* = \min_i\{\mathrm{EV}_i\}$, 分别对应于问题的益损值是代表收益 (收入) 或是损失 (支出).

习题 11.2A

1. 你受邀请参加电视上的幸运转轮游戏. 这种电动转轮有两个按钮, 分别用于启动轮子的硬转动 (H) 或软转动 (S). 转轮本身被划分成白色 (W) 和红色 (R) 各半圈区域. 并且告诉你, 轮子停在白色区域的概率为 0.3, 停在红色区域的概率是 0.7. 你参与该游戏的益损值为

	W	R
H	\$800	\$200
S	−\$2 500	\$1 000

画出相应的决策树并确定行动方案.

***2.** McCoy 是位农场主, 他正在考虑种玉米还是种大豆. 这两种商品作物明年收获期的价格是涨价、平价、还是跌价的概率分别为 0.25, 0.30, 0.45. 假如价格上涨, 种玉米的净收入为 \$30 000, 而种大豆的纯收益是 \$10 000. 如果价格不变的话, McCoy 将会 (刚好) 不赔不赚. 而一旦价格下跌, 玉米和大豆将分别损失 \$35 000 和 \$5 000.

(a) 用决策树表示 McCoy 的问题. (b) McCoy 该种哪种作物?

3. 你有机会对 3 种公共基金进行投资: 公共事业基金 (U)、进取型增长基金 (A) 和全球基金 (G). 你的投资价值会随着市场情况而变化. 有 10% 的机会市场会下滑, 有 50% 的机会保持平稳, 而有 40% 的机会市场表现良好. 下表提供了在这 3 种情况下市场变化的百分比:

可能的方案	投资回报百分比		
	市场下滑 (%)	市场平稳 (%)	市场上扬 (%)
U	+5	+7	+8
A	−10	+5	+30
G	+2	+7	+20

(a)用决策树表示该问题.　　　　(b)你该选择种哪种基金投资?

4. 你可以投资一种利息为 7.5% 的债券, 能按券面价值卖出; 也可以投资一种进取型增长股票, 仅有 1% 的红利. 如果担心通货膨胀, 利率就会上升到 8%, 在这种情况下债券的本金价值会下降 10%, 股票价值会降低 20%. 假如真的出现经济不景气, 利率将会下降 6%, 在这种条件下, 债券本金值预计上升 5%, 而股票价值会增加 20%. 如果经济保持不变, 该只股票的价值会上升 8%, 债券本金价值保持不变. 有经济学家预测, 出现通货膨胀的机会有 20%, 经济衰退的的机会有 15%. 假设你正在根据明年的经济条件作出投资决策.
(a) 用决策树表示这个问题.　　　　(b) 你会投资股票还是投资债券?

5. AFC 即将在全国推出新的"鸡翅套餐"快餐产品. 公司研究部门相信, "鸡翅套餐"将会非常成功, 并打算不做广告而直接将该产品立即在所有 AFC 连锁店推出. 但市场部门对"套餐"的意见不同, 想要进行大力度的广告宣传. 这项广告宣传将耗费 \$100 000, 如果成功则会产生 \$950 000 的收益. 假如广告不成功 (不成功的机会有 30%), 则预计收益只有 \$200 000. 如果不做广告而得到顾客接受, 预计的收益为 \$400 000, 其概率为 0.8; 如果顾客不接受, 则收益只有 \$200 000, 概率为 0.2.
(a) 画出相应的决策树.　　　　(b) AFC 应该采取哪种行动来推出这个新产品?

*6. 一枚硬币被连续抛 3 次, 每正面 (H) 朝上 1 次你会得到 \$1.00, 每接连 2 次正面朝上你还会另外得到 \$0.25(提醒, HHH 包括 2 组 HH). 但是, 每次反面朝上时, 你就需要退回 \$1.1. 你可以选择玩还是不玩这个游戏.
(a) 画出该游戏的决策树.　　　　(b) 玩这个游戏对你有利吗?

7. 你有个机会在赌场玩下面的游戏. 一个骰子连掷 2 次, 会出现 4 种结果: (1) 2 次都是偶数点; (2) 2 次都是奇数点; (3) 2 次结果是偶数–奇数点, 或者是奇数–偶数点; (4) 所有的其他情况. 赌场允许你只赌 2 种结果, 赌金相同. 例如, 你可以下注同样的钱数赌偶数点 (结果 1) 和奇数点 (结果 2). 你赌的每元钱的收益是: 赌中第 1 种结果得 2 元钱, 赌中第 2 种和第 3 种结果得 1.95 元, 赌中第 4 种结果得 1.50 元.
(a) 画出该游戏的决策树.　　　　(b) 你该作出哪两种选择?
(c) 在这场游戏中, 你能赢钱吗?

8. Acme 制造公司批量生产某种装饰品, 产品不合格率为 0.8%、1%、1.2% 和 1.4% 的相应概率分别为 0.4, 0.3, 0.25, 0.05. 有 3 个客户 A, B, C 要签订批量订货合同, 分别要求不合格率为 0.8%、1.2% 和 1.4%. 假如产品不合格率高于合同规定的, 每增加 0.1% 就会对 Acme 公司罚款 \$100. 反过来, 供应比合同要求更高质量的产品批次的话, 每提高 0.1% 则 Acme 会增加成本 \$50. 假定在送货前并不对批量产品进行检查.
(a) 画出相应的决策树.　　　　(b) 这 3 个客户中哪个应该优先得到他们的订单?

9. TriStar 公司计划在阿拉斯加开设一家新的工厂. 该公司可选择现在就开设一家完整规模的大工厂, 也可以选择先开设一家小工厂, 等 2 年后需求条件改善时再扩大. 这个决策问题要考虑的时间范围为 10 年, TriStar 预计未来 10 年高需求和低需求的概率分别为 0.75 和 0.25. 现在建大厂的费用是 500 万美元, 而建一个小厂的费用为 100 万美元. 2 年后扩建一个小厂的费用为 420 万美元. 下表给出未来 10 年运作的收入情况:

可能的方案	预计年收入 (单位 $1 000)	
	高需求	低需求
现在建大厂	1 000	300
现在建小厂	250	200
两年后扩建	900	200

(a) 做出相应的决策树, 已知 TriStar 可以选择两年以后扩建或不扩建小厂.

(b) 为 TriStar 公司制订一个未来 10 年的建设策略.(为简单起见, 不考虑货币的时间价值.)

10. 重新求解第 9 题, 假定年利率为 10%, 并在决策中考虑货币的时间价值. [注意: 在求解这个问题中你需要计算复利. 可以用 Excel 函数 NPV(i, R) 对利率 i 计算储存在范围 R 上的现金流价值. NPV 假设每笔现金流都发生在年底.]

11. 重新求解第 9 题, 假设需求可以分为高、中、低, 相应的概率分别为 0.7, 0.2, 0.1. 只有当头 2 年的需求为高的情况下才扩建小厂. 下表给出年收入, 不计货币的时间价值.

可能的方案	预计年收入 (单位 $1 000)		
	高需求	中需求	低需求
现在建大厂	1 000	500	300
现在建小厂	400	280	150
2 年后扩建	900	600	200

12. Sunray 电器公司有一个 200 辆卡车的车队为它的电器网络提供服务, 该公司想要为车队制定一个预防性的定期维修计划. 车队第 1 年的故障概率为零, 第 2 年的故障概率为 0.03, 从第 3 年到第 10 年每年增加 0.01. 超过 10 年后, 故障率是个常数, 为 0.13. 每辆卡车一次突发故障的维修费为 $200, 且每辆卡车的定期维修费只有 $75. Sunray 公司希望为这些定期的预防性维修确定最优周期 (按月).

(a) 建立相应的决策树. (b) 确定最优的维修周期长度.

13. 某家食品店对面包的日需求量按下面的概率分布确定:

n	100	150	200	250	300
p_n	0.20	0.25	0.30	0.15	0.10

该店以 55 美分/块的价格买进面包, 再按照 $1.2/块的价格售出. 每天卖不掉的面包按照 25 美分/块处理. 假定库存水平由 p_n 指定的某个需求水平决定.

(a) 建立相应的决策树. (b) 该店每天应存多少块面包?

14. 在第 13 题中, 假设该店想要把这个决策问题扩充为 2 天期间的情形, 第 2 天方案取决于第 1 天的需求情况. 如果第 1 天的需求正好等于库存量, 该店就会按第 1 天同样的量订货. 假如需求超过库存量的话, 该店就会多订, 而如果少于库存的话, 就会少订. 建立相应的决策树并确定出最优订货策略.

***15.** 某自动机器厂生产 α(单位: 1 000) 件产品. 随着 α 增加, 其次品率 p 随以下概率密度函数增加:

$$f(p) = \begin{cases} \alpha p^{\alpha-1}, & 0 \leqslant p \leqslant 1 \\ 0, & \text{其他} \end{cases}$$

如出现 1 件次品, 公司损失 \$50. 对每件合格产品, 公司将获利 \$5.

(a) 建立相应的决策树.　**(b)** 求出使公司获得最大期望利润的 α 值.

16. 在一台机床上加工外径为 d 的一种圆柱体, 上下误差允许范围为 $d+t_U$ 和 $d-t_L$. 生产工艺服从正态分布, 平均值为 μ, 标准差为 σ. 直径过大的圆柱体需要返工, 返工费用为 c_1 元. 如果圆柱体过小就成了废料, 费用为 c_2 元. 建立决策树并求最优设定的 d.

17. (Cohan 等人, 1984) 现代林业管理采用受控山火的方法来降低火灾危险并刺激新林木的生长. 管理部门可选择延期点火或按计划点火. 在某个具体的开阔地上, 如果延期点火, 将会引致 \$300 的管理费. 如果按计划实施控制点火, 则有 50% 的机会是好天气, 并且实施燃烧的费用为 \$3 200. 燃烧的结果可能成功, 成功的概率为 0.06; 也可能燃烧得不成功, 概率是 0.4. 实施成功的话, 预计产生效益 \$6 000; 实施不好时效益只有 \$3 000. 假如天气不好, 将取消燃烧, 相应的计划成本是 \$1 200, 不产生任何效益.

(a) 建立一个决策树来确定燃烧是该按计划进行还是该延期.

(b) 研究所求的解对好天气概率变化的灵敏度.

18. (Rappaport, 1967) 某制造厂采用了线性规划方法确定所生产的各种电视型号的最优生产量组合. 制造厂最近得到的信息显示, 某个型号所用的零部件的供应商有 40% 的机会涨价 \$35. 由此制造厂可以采取两种行动: 继续使用原来的 (最优) 产品组合 (A1), 或者根据提高了的零部件价格采用新的 (最优) 组合 (A2). 假如价格不涨的话, 行动 A1 最理想; 如果涨价的话, 行动 A2 最理想. 下表给出了各方案关于零部件价格的每月总利润, 它随着所采取的行动相应产生的随机性后果而变化:

	价格上涨 (O1)	价格不涨 (O2)
原来的产品组合 (A1)	\$400 000	\$295 500
新的产品组合 (A2)	\$372 000	\$350 000

(a) 建立相关的决策树并决定应采取的行动.

(b) 该制造厂可投资 \$1 000 以得到是否涨价的额外信息. 从这一信息可知道, 涨价的概率为 0.9 的机会有 58%, 涨价概率为 0.3 的机会有 42%. 你会建议投入这笔额外的资金吗?

***19.** **渴望水平指标.** Acme 制造公司在其生产工艺中利用一种工业化学品. 该化学品的有效期为 1 个月, 过期没用的必须销毁. Acme 公司对这种化学品的用量 (单位: 加仑) 是随机的, 服从以下分布:

$$f(x)=\begin{cases}\dfrac{200}{x^2}, & 100\leqslant x\leqslant 200\\ 0, & \text{其他}\end{cases}$$

该化学品的实际消费量在月初立即出现. Acme 想要确定化学品的库存量, 满足两个冲突的指标 (也称渴望水平指标): 月度平均剩余量不超过 20 加仑, 并且月度平均缺货量不超过 40 加仑.

11.2.2 期望值指标的各种变化

本节解决有关期望值指标的 3 个问题. 第 1 个问题是根据实验结果确定*后验概率*, 第 2 个问题涉及*效用*指标与货币实际价值指标的比较.

后验 (贝叶斯) 概率 用于期望值指标的概率通常从历史数据得出 (见 14.5 节). 在某些情况下, 这些概率能通过一些根据抽样和实验方法得出的其他信息做出改进, 这样产生的概率就称为**后验 (或称贝叶斯) 概率** [Posterior (或 Bayes') Probability], 它和从原始数据得出的**先验概率** (prior probability) 正好相反.

实际应用——Casey的问题：全新化验结果的解释和评估

有个新生儿名叫 Casey, 对她的一次初筛化验结果显示出她缺少一种 C14:1 酶, 人需要这种酶来消化一种特殊形式的脂肪, 一旦缺少这种酶就会产生严重的疾病, 也可能导致原因不明的死亡 (一般属于婴儿突然死亡综合症, 简称 SIDS). 以前对大约 13 000 个新生儿做过这项化验, Casey 是第一例呈阳性的结果. 尽管只凭这项化验还不能做出确切的诊断, 但这种情况出现的极端稀少性还是让她的医生相信, Casey 患上了这一缺乏症的可能性有 80%~90%. 由于 Casey 的化验呈阳性, 医院用了贝叶斯后验概率来评估这个孩子是否真的患有 C14:1 缺乏症. 第 24 章中的案例 11 给出了对这个问题的详细分析.

例 11.2-2

本例说明如何对期望值指标进行修改, 以便能利用后验概率信息. 在例 11.2-1 中, 从已有的金融出版物上, 我们分别得到了牛市和熊市的 (先验) 概率为 0.6 和 0.4. 假设你不仅仅是依赖于这些出版物信息, 而是决定通过咨询一位在股票市场方面做得很好的朋友, 做一次更加“个性化”的调查. 这位朋友提供了“赞成”或“反对”投资的一般性观点, 我们把它按如下方式进行了量化: 如果是牛市, 则投“赞成”票的机会有 90%; 但如果是熊市的话, 投“赞成”的机会只有 50%. 你如何利用这些额外的信息呢?

朋友的这段叙述提供了已知牛市和熊市的自然状态下的“赞成/反对”的条件概率. 为便于表达, 我们采用了下列符号:

$$v_1 = \text{投“赞成”票}$$
$$v_2 = \text{投“反对”票}$$
$$m_1 = \text{熊市}$$
$$m_2 = \text{牛市}$$

这位朋友的叙述可以写成以下概率公式:

$$\mathrm{P}\{v_1|m_1\}=0.9,\quad \mathrm{P}\{v_2|m_1\}=0.1$$
$$\mathrm{P}\{v_1|m_2\}=0.5,\quad \mathrm{P}\{v_2|m_2\}=0.5$$

有了这些额外的信息, 这一决策问题可总结如下:

(1) 如果朋友的建议是“赞成”, 你会投资股票 A 还是股票 B?

(2) 如果朋友的建议是“反对”, 你会投资股票 A 还是股票 B?

这个问题可以用图 11.5 所示的决策树形式. 节点 1 是一个机会事件, 代表“赞成”和“反对”的概率. 节点 2 和 3 是根据朋友对“赞成”和“反对”的各自投票结果, 而在股票 A 和股票 B 之间选择的决策点. 最后, 节点 4 至节点 7 是机会事件, 代表牛市和熊市.

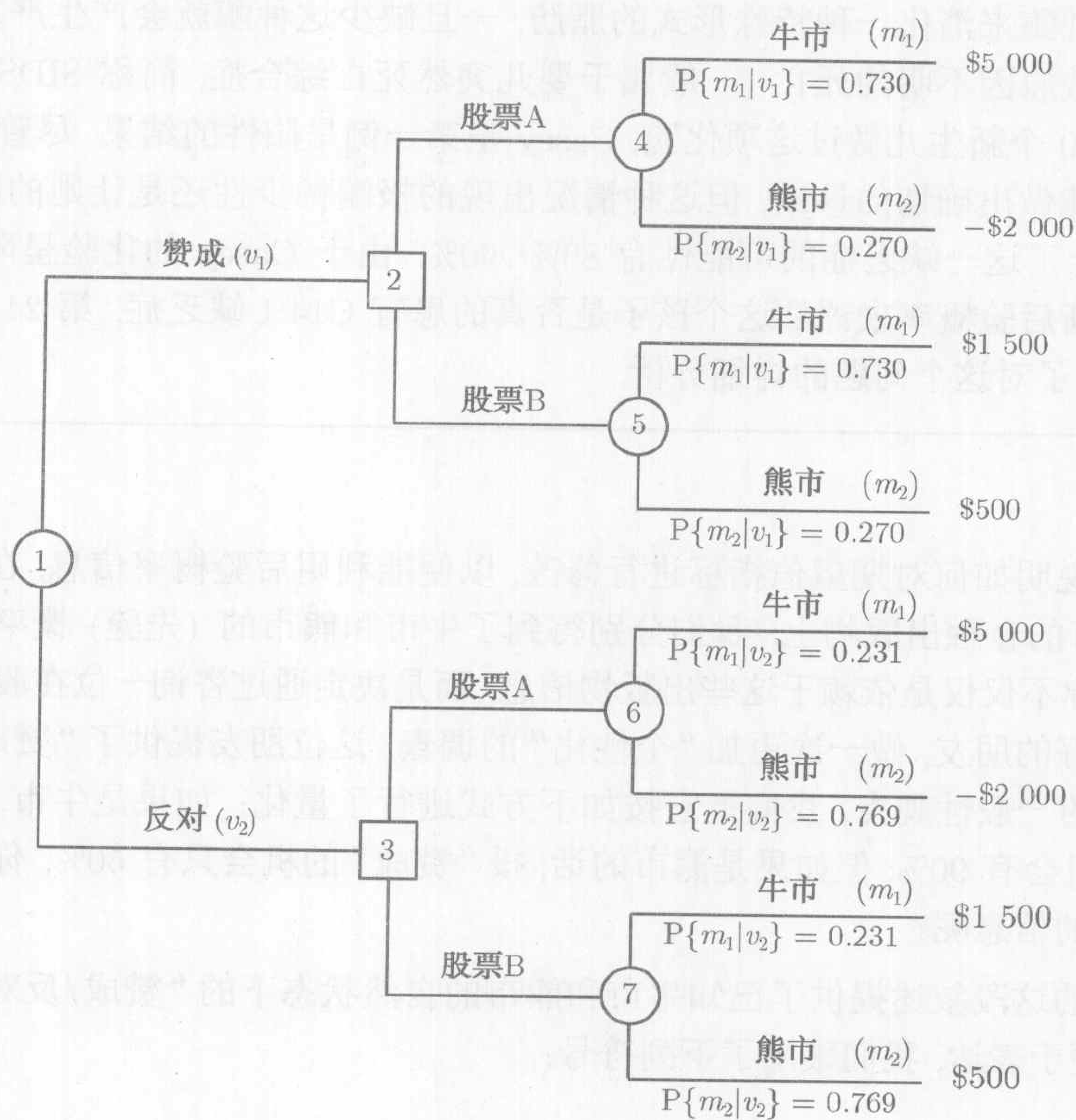


图 11.5　带有后验概率的股票市场问题决策树

为了评估图 11.5 的不同选择方案, 必须要计算机会节点 4, 5, 6, 7 的 m_1 和 m_2 分支的后验概率 $\mathrm{P}\{m_i|v_j\}$, 这些后验概率提供了朋友的“赞成”和“反对”建议的额外信息, 根据下列一般步骤计算:

第 1 步 该问题的条件概率 $P\{v_j|m_i\}$ 为

	v_1	v_2
m_1	0.9	0.1
m_2	0.5	0.5

第 2 步 计算各联合概率

$$P\{m_i, v_j\} = P\{v_j|m_i\}P\{m_i\}, \quad \text{对所有的}i \text{ 和 } j$$

由先验概率 $P\{m_1\} = 0.6, P\{m_2\} = 0.4$, 求上述联合概率只要将第 1 步表中的第 1 行和第 2 行分别用 0.6 和 0.4 相乘, 我们得到

	v_1	v_2
m_1	0.54	0.06
m_2	0.20	0.20

表中各项之和等于 1.

第 3 步 计算绝对概率

$$P\{v_j\} = \sum_{\text{所有}i} P\{m_i, v_j\}, \quad \text{对所有的}j$$

这些概率值可从第 2 步表中每行相加算出

$P\{v_1\}$	$P\{v_2\}$
0.74	0.26

第 4 步 计算所需的后验概率

$$P\{m_i|v_j\} = \frac{P\{m_i, v_j\}}{P\{v_j\}}$$

为计算这些概率值, 将第 2 步表中每列各行用第 3 步表中对应列元素相除, 得到

	v_1	v_2
m_1	0.730	0.231
m_2	0.270	0.769

这些就是图 11.5 中所显示的概率, 与原来的先验概率 $P\{m_1\}$=0.6 和 $P\{m_2\}$ = 0.4 不同.

现在, 我们就可以根据节点 4 至节点 7 的期望益损值来评估各方案了, 即

"赞成"

节点 4 的股票 A = 5 000 × 0.730 + (−2 000) × 0.270 = $**3 110**

节点 5 的股票 B = 1 500 × 0.730 + 500 × 0.270 = $1 230

决策：投资股票 A.

“反对”

$$\text{节点 6 的股票 A} = 5\ 000 \times 0.231 + (-2\ 000) \times 0.769 = -\$383$$

$$\text{节点 7 的股票 B} = 1\ 500 \times 0.231 + 500 \times 0.769 = \$\mathbf{731}$$

决策：投资股票 B.

上面的决策相当于说，在决策点 2 和决策点 3 的期望收益分别为 \$3 110 和 \$731(见图 11.5). 因此，对于第 3 步算出的概率 $P\{v_1\}$ 和 $P\{v_2\}$, 我们可以计算出整个决策树的期望收益 (见习题 11.2B 的第 3 题).

Excel 程序

Excel 文件 excelBayes.xls 用来对最大 10×10 的先验概率矩阵 (行 7:14 和列 F:K 及 O:V 被隐藏起来以保留空间) 计算贝叶斯后验概率. 输入数据包括 $P\{m\}$ 和 $P\{v|m\}$. 该电子表格程序检查输入数据中的错误，并显示适当的错误信息. 图 11.6 显示了对例 11.2-2 问题模型的应用.

	A	B	C	D	E	L	M	N
1	Bayes Posterior Probabilities							
2	Input Data						Output Results	
3			P{v|m} (10 x 10) maximum					P{v,m}
4	P{m}		v1	v2			v1	v2
5	m1	0.6	0.9	0.1			0.5400	0.0600
6	m2	0.4	0.5	0.5			0.2000	0.2000
15	Input Data Error Messages						P{v}	
16							0.7400	0.2600
17							P{m|v}	
18						m1	0.7297	0.2308
19						m2	0.2703	0.7692
20								
21								
22								
23								
24								
25								
26								
27								

图 11.6　用 Excel 计算例 11.2-2 的贝叶斯后验概率 (文件 excelBayes.xls)

习题 11.2B

1. 某社区学院的数据显示，和没有学过微积分的 50% 的学生相比，在高中学过微积分的新生中有 75% 学习成绩都不错. 当年的招生情况表明，有 30% 的新生已经学过微积分. 那么一名新生在学院里学习成绩好的概率是多少？

***2.** Elektra 公司有 75% 的电器元件是从供应商 A 采购的, 剩下的 25% 来自供应商 B. 供应商 A 和 B 产品的次品率分别是 1% 和 2%. 在收到一批货中随机抽样 5 件进行检查时, 只发现 1 件不合格品. 求该批量产品是来自供应商 A 的概率. 对供应商 B 重复同样的计算.(提示: 一个样本中不合格品的概率服从二项分布.)

3. 在例 11.2-2 中, 假定你有另外一项投资选择, 把初始的 \$10 000 投入到某个安全的定期储蓄中, 储蓄利率为 8%. 你朋友建议只用于股票投资.

(a) 建立相应的决策树.

(b) 这种情况下最优的决策是什么?(提示: 利用例 11.2-2 第 3 步给出的 $P\{v_1\}$ 和 $P\{v_2\}$, 计算出投资股票市场的期望值.)

***4.** 你是一名有希望写出一本成功小说的作家. 你可以选择自己出版这本小说, 也可以通过出版社来出版. 如果签了合同, 出版社给你 \$20 000. 如果小说成功, 会有 200 000 本的销售量; 假如不成功的话, 只能卖出 10 000 本. 出版社支付每本版税 \$1. 出版社进行的一项市场调查表明, 这本小说成功的机会有 70%. 如果你自己出版这本小说, 一开始你要支付 \$90 000 用作印刷和宣传, 但卖出每本书可净赚 \$2.

(a) 根据以上信息, 你会接受出版商的条件, 还是自己出版这本小说?

(b) 假设你委托一个文学代理公司进行一项小说是否可能成功的调查. 根据过去的经验, 该代理公司告诉你说, 当小说成功情况下, 有 20% 的时候调查预测出错误的结果; 而当小说不成功的情况下, 有 85% 的时候调查给出正确的预测. 这个信息对你的决策有怎样的影响?

5. 考虑习题 11.2A 第 2 题中农场主 McCoy 的决策问题. 假设这位农场主还有另外一种选择, 即把土地用作牧场, 在这种情况下保证能有 \$7 500 的收益. 这位农场主还从经纪人那里得到了有关未来商品价格稳定性的确切信息. 经纪人对这一方案“有利”和“不利”的评价进一步量化为下面的条件概率:

$P\{a_j|s_i\} =$

	a_1	a_2
s_1	0.85	0.15
s_2	0.50	0.50
s_3	0.15	0.85

符号 a_1 和 a_2 表示经纪人“有利”和“不利”的评估结果, s_1, s_2, s_3 分别代表未来价格的升、平、降变化.

(a) 建立相应的决策树. (b) 确定该问题的最优决策.

6. 习题 11.2A 第 5 题中, 假设 AFC 的管理部门决定在指定地区对它的鸡翅套餐进行市场检验. 检验的结果要么是“好”, 要么是“坏”. 这项检验产生了下列条件概率, 分别对应于是否进行广告宣传:

$P\{a_j|v_i\}$——作广告

	a_1	a_2
v_1	0.95	0.05
v_2	0.3	0.7

$P\{a_j|w_i\}$——不作广告

	a_1	a_2
w_1	0.8	0.2
w_2	0.4	0.6

符号 v_1 和 v_2 代表“成功”和“不成功”, w_1 和 w_2 代表“可以接受”和“不能接受”.

(a) 建立相应的决策树.　　(b) 确定出 AFC 的最佳行动方案.

7. 根据 Acme 制造公司的历史数据估算, 所生产出的小饰品批量不能令人接受 (坏的) 的机会有 5%. 一次坏批量中有 15% 的次品, 而一次好批量只有 4% 的不合格品. 令 $a = \theta_1 (= \theta_2)$ 表示该批量是好的 (坏的), 相应的先验概率为

$$\mathrm{P}\{a = \theta_1\} = 0.95, \quad \mathrm{P}\{a = \theta_2\} = 0.05$$

不是只根据先验概率发货, 公司还做了两件产品的抽样检验, 得出 3 种可能的结果: (1) 这两件都是好的 (z_1), (2) 有一件是好的 (z_2), (3) 这两件都是次品 (z_3).

(a) 求出后验概率 $\mathrm{P}\{\theta_i | z_j\}, i = 1, 2; j = 1, 2, 3$.

(b) 假设该制造公司把各批量产品供货给客户 A 和客户 B, 而且合同规定, 客户 A 和 B 的次品率分别不超过 5% 和 8%, 超过以上最大上限的每个百分点将支付 \$100 罚金. 供应比合同规定更好质量的产品时, 每多一个百分点公司要多花费 \$50. 建立相应的决策树并确定供货批量的优先策略.

效用函数　在前面介绍的内容中, 期望值指标已被应用在收益是实际 货币的情况下. 然而还有一些情况下, 分析中应该用到的是效用, 而不是实际的货币价值. 为了说明这一点, 假设有 50:50 的机会, 使得一笔 \$20 000 的投资要么产生 \$40 000 的收益, 要么完全损失掉, 相应的期望收益为 $40\,000 \times 0.5 - 20\,000 \times 0.5 = \$10\,000$. 虽然存在一笔净收益, 但不同的人对这个结果的解释却大相径庭. 一个愿意接受风险的投资人可能会因为有 50% 的机会赚 \$40 000 而进行投资; 反过来, 一个保守的投资人则可能不愿意冒损失 \$20 000 的风险. 从这个观点来看, 不同的人面对风险显示出不同的态度, 也就是说, 个人关于风险显示出不同的效用.

效用的确定是主观性的, 依赖于我们对接受风险的态度. 本节介绍一种关于决策者对风险承受能力的量化方法. 最终的结果是用一种效用函数来代替货币价值.

在上述投资情形中, 最好的益损值是 \$40 000, 最差益损值是 $-\$20\,000$. 这样我们建立一种任意 (但合理) 的效用度量 U, 从 0 到 100, 其中 $U(-\$20\,000) = 0, U(\$40\,000) = 100$. 对 $-\$20\,000$ 到 \$40 000 中间值的效用按以下方法确定: 如果决策者对风险的态度一般, 则产生的效用函数为连接 $(0, -\$20\,000)$ 到 $(100, \$40\,000)$ 的一条直线. 在这种情况下, 真正的货币值和它的效用都会产生同样的决策. 实际上, 效用函数采用了其他形式来反映决策者对风险的态度. 图 11.7 显示了 X, Y, Z 这 3 个人的情况. X 属于**保守型**(小心谨慎型) 的人, 因为图中显示出他对损失比收益更加敏感. Z 与之正相反, 因此他**勇于冒险**. 图中显示, 对于保守型的 X, 对应于损失 $-\$10\,000$ 的效用的降幅 bc 比相应于获利 \$10 000 的增幅 ab 来得大. 对同样的 $\pm\$10\,000$ 变化, 勇于冒险的 Z 显现出完全相反的行为, 因为 $de > ef$. 此外, Y 属于**稳健型**的人, 因为对应于上述变化的效用是相同的. 一般地, 某个人可能既是保守型的又同时具有是冒险型的一面, 在这种情况下, 相应的效用曲线显现出某种拉长的 S 形.

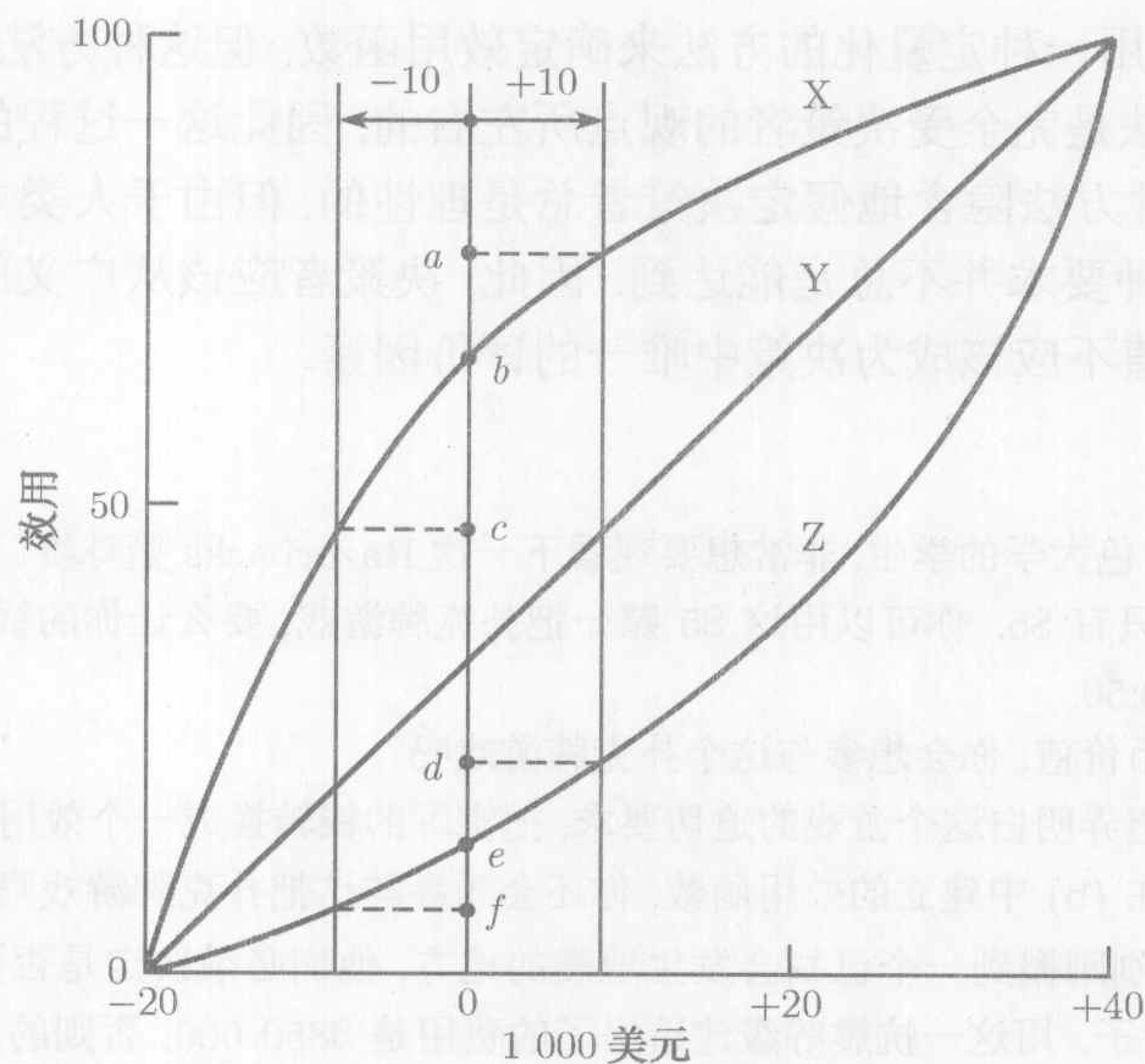


图 11.7 保守型 (X)、稳健型 (Y) 和冒险型 (Z) 决策者的效用函数

类似于图 11.7 上显示的效用曲线是通过将决策者对不同现金货币水平风险的态度进行"量化"所得到的. 在我们的例子中, 所需的区间为 (−\$20 000, \$40 000), 相应的效用值范围是 (0, 100). 我们想做的是, 对所有中间的货币值, 如 −\$10 000、\$0、\$10 000、\$20 000 和 \$30 000, 给出其相应的效用值. 具体方法从给定一个现金量 x 建立一次博彩开始, 其期望效用函数为:

$$\begin{aligned}U(x) &= pU(-20\,000) + (1-p)U(40\,000), 0 \leqslant p \leqslant 1\\ &= 0p + 100(1-p)\\ &= 100 - 100p\end{aligned}$$

为了确定 $U(x)$, 我们让决策者说出是愿意保有现金值 x, 还是愿意利用这个机会博一次彩, 其中输掉 \$20 000 的概率是 p, 而赢得 \$40 000 的概率为 $1-p$. 决策者把他的偏好转述为指定一个 p 的值, 这个值使得他对这两个选择都无差别. 比如, 假设 $x = \$20\,000$, 决策者可能就会说, 当 $p = 0.8$ 时, 保有这 \$20 000 和博彩的吸引力是等同的. 在这种情况下, 我们可计算出 $x = \$20\,000$ 的效用为

$$U(20\,000) = 100 - 100 \times 0.8 = 20$$

我们不断采用这种方式, 直到产生出足够的点 $[x, U(x)]$ 来得到效用函数的形状. 然后我们就可以利用回归分析或者分段线性函数, 来得出我们所需要的效用函数了.

尽管我们在用一种定量化的方法来确定效用函数, 但这种方法非常不科学. 事实上, 这样的方法是完全受决策者的观点所左右的, 因此这一过程的可靠性令人怀疑. 特别是, 这种方法隐含地假定决策者总是理性的, 但由于人类本身的行为和情绪的变化性, 这种要求并不总是能达到. 因此, 决策者应该从广义的角度理解效用的概念, 货币价值不应该成为决策中唯一的评价因素.

习题 11.2C

*1. 你是一名阿肯色大学的学生, 非常想要观看下一次 Razorbacks 篮球赛, 但问题是买门票需要 $10, 而你只有 $5. 你可以用这 $5 赌一把扑克牌游戏, 要么让你的钱翻一番, 要么全输掉, 机会是 50:50.

(a) 按照货币价值, 你会想参与这个扑克牌游戏吗?

(b) 根据你想弄明白这个游戏的迫切要求, 把实际的钱转换成一个效用函数.

(c) 按照你在 (b) 中建立的效用函数, 你还会想参与这把扑克牌游戏吗?

*2. Golden 一家刚刚搬到一个已知会发生地震的地方, 他们必须决定是否要按照高抗震等级建造自己的房子. 用这一抗震等级建造房子的费用是 $850 000, 否则的话造一座同类房子的价格只有 $350 000. 假如地震发生了 (而且可能发生地震的概率有 0.001), 一座抗震等级不够的房子就会花 $900 000 去修缮. 建立与该问题相关的博彩, 假定效用度量为 0 到 100.

*3. 在一项高风险下的冒险投资中, 一笔 $10 000 在下一年内有 50:50 的机会要么增值到 $14 000, 要么贬值剩 $8 000. 这样, 净收益可能是 $4 000, 也可能是 −$2 000.

(a) 假设投资人对风险不敏感, 效用度量从 0 到 100, 确定投资净收益为 $0 的效用, 以及相应的无差别概率.

(b) 假设有两个投资人 A 和 B 显示出下列无差别概率:

净收益 ($)	无差别概率	
	投资人 A	投资人 B
−2 000	1.00	1.00
−1 000	0.30	0.90
0	0.20	0.80
1 000	0.15	0.70
2 000	0.10	0.50
3 000	0.05	0.40
4 000	0.00	0.00

画出投资人 A 和投资人 B 的效用函数曲线, 并把每个投资人划定为是不愿冒险的投资人还是勇于冒险的投资人.

(c) 假定投资人 A 有机会对两个项目 I 或 II 之一进行投资. 项目 I 可能产生 $3 000 的净收益, 概率是 0.4; 也可能净损失 $1 000, 概率为 0.6. 项目 II 有可能产生 $2 000 的净收益, 概率是 0.6; 也可能没收益, 概率为 0.4. 根据 (b) 中的效用函数, 利用期望效用指标来确定投资人 A 应选哪一项目投资. 对应所选项目的期望费用值是多少?(提示: 利用该效用函数的线性插值.)

(d) 对投资人 B 重复 (c).

11.3 不确定型决策

和风险型决策一样, 不确定型决策也涉及其益损依赖于 (随机)*自然状态* 的行动方案. 具体来说, 一个有 m 个行动方案和 n 个自然状态的决策问题的益损矩阵可以表示为

$$
\begin{array}{c|cccc|}
 & s_1 & s_2 & \cdots & s_n \\
a_1 & v(a_1, s_1) & v(a_1, s_2) & \cdots & v(a_1, s_n) \\
a_2 & v(a_2, s_1) & v(a_2, s_2) & \cdots & v(a_2, s_n) \\
\vdots & \vdots & \vdots & & \vdots \\
a_m & v(a_m, s_1) & v(a_m, s_2) & \cdots & v(a_m, s_n)
\end{array}
$$

元素 a_i 代表行动 i, 元素 s_j 代表自然状态 j. 对应于行动 a_i 和自然状态 s_j 结果的益损值为 $v(a_i, s_j)$.

风险型决策和不确定型决策的区别是, 在不确定型决策情况下, 对应于状态 $s_j, j = 1, 2, \cdots, n$ 的概率分布是未知的或无法确定. 这种信息的缺乏决定了需要建立以下准则来分析这种决策问题:

(1) 拉普拉斯准则; (2) 极小化极大准则;

(3) 损失后悔准则; (4) 赫尔维茨准则.

这些准则的差别在于决策者面对不确定情形下的保守程度.

拉普拉斯准则建立在**不充分理由原则** (principle of insufficient reason) 之上. 由于概率分布未知, 因此没有理由相信有关自然状态的概率是不同的. 因此对可能方案的评估就采用了*乐观性* 的假设, 即所有的状态都等可能发生, 即 $\mathrm{P}\{s_1\} = \mathrm{P}\{s_2\} = \cdots = \mathrm{P}\{s_n\} = \frac{1}{n}$. 设益损值 $v(a_i, s_j)$ 代表收益, 则最好的那个方案将达到

$$\max_{a_i}\left\{\frac{1}{n}\sum_{j=1}^{n} v(a_i, s_j)\right\}$$

假如 $v(a_i, s_j)$ 表示损失, 则用最小化代替最大化.

极大化极小 (极小化极大)准则基于让最差可能条件达到最好的保守态度. 如果 $v(a_i, s_j)$ 为损失, 则我们选择对应于**极小化极大准则**

$$\min_{a_i}\left\{\max_{s_j} v(a_i, s_j)\right\}$$

的行动方案. 假如 $v(a_i, s_j)$ 是收益, 我们采用下面给出的**极大化极小准则**

$$\max_{a_i}\left\{\min_{s_j} v(a_i, s_j)\right\}$$

损失后悔准则寻求在极小化极大(极大化极小)指标中缓和保守主义，用一种*损失 (或称为后悔)*矩阵 $r(a_i, s_j)$ 通过下面的变换来代替 (收益或损失的) 益损矩阵 $v(a_i, s_j)$:

$$r(a_i, s_j) = \begin{cases} v(a_i, s_j) - \min\limits_{a_k} \{v(a_k, s_j)\}, & \text{如果 } v \text{ 为损失} \\ \max\limits_{a_k} \{v(a_k, s_j)\} - v(a_i, s_j), & \text{如果 } v \text{ 为收益} \end{cases}$$

为了说明为什么损失后悔准则“缓和了”极小化极大 (极大化极小) 准则，我们考虑下面的损失矩阵 $v(a_i, s_j)$

	s_1	s_2	行最大值
a_1	\$11 000	\$90	\$11 000
a_2	\$10 000	\$10 000	**\$10 000** ← 极小化极大值

使用极小化极大准则，那么最好采用具有有限损失 \$10 000 的 a_2. 但是我们也可以选择 a_1, 因为如果只有 s_2 出现时，我们有机会把损失限制在 \$90.

我们来看看假如使用下面的后悔矩阵 $r(a_i, s_j)$ 时情况会怎样：

	s_1	s_2	行最大值
a_1	\$1 000	\$0	**\$1 000** ← 极小化极大值
a_2	\$0	\$9 910	\$9 910

用这个后悔矩阵时，该极小化极大准则将会根据需要选择 a_1.

最后来看**赫尔维茨**准则，它用来反映从最乐观到最悲观 (或保守) 的决策态度. 定义 $0 \leqslant \alpha \leqslant 1$, 并假设 $v(a_i, s_j)$ 表示收益，则选择的行动必须与下式相关：

$$\max_{a_i} \left\{ \alpha \max_{s_j} v(a_i, s_j) + (1-\alpha) \min_{s_j} v(a_i, s_j) \right\}$$

α 称为**乐观指数** (index of optimism). 如果 $\alpha = 0$, 该指标是保守的，因为它用了通常的极小化极大准则. 如果 $\alpha = 1$, 则该指标产生乐观的结果，因为它寻求好中最好的条件. 我们可以通过在 $(0, 1)$ 区间内选择适当的 α 值，来对乐观 (或悲观) 程度作一些调整. 在对乐观性和悲观性感觉同样强的时候，我们可以适当选择 $\alpha = 0.5$.

如果 $v(a_i, s_j)$ 表示损失，则该准则改变为

$$\min_{a_i} \left\{ \alpha \min_{s_j} v(a_i, s_j) + (1-\alpha) \max_{s_j} v(a_i, s_j) \right\}$$

例 11.3-1

全国户外运动学校 (National Outdoors School, NOS) 正准备在阿拉斯加中心地区建一个夏季宿营地，用来作为培训野外生存人员的基地. NOS 估计，参加培训的人数可分为 4 档：200, 250, 300 和 350. 当规模正好符合需求时，建营地的费用最小. 如果超过和不足需求水平时，或者由于建造多余的 (用不了的) 营地，或者由于

营地建造不足、满足不了需求而损失收入机会, 都会发生额外的费用. 令 a_1 到 a_4 表示营地规模 (200, 250, 300, 350), s_1 到 s_4 是参加人员数, 下表给出了该情形的费用矩阵 (单位为 1 000 美元).

	s_1	s_2	s_3	s_4
a_1	5	10	18	25
a_2	8	7	12	23
a_3	21	18	12	21
a_4	30	22	19	15

用上述 4 种准则分别分析这个问题.

拉普拉斯准则 设 $\mathrm{P}\{s_j\}=\frac{1}{4}, j=1,2,3,4$, 不同行动方案的期望值计算为

$$\mathrm{E}\{a_1\}=\tfrac{1}{4}(5+10+18+25)=\$14\,500$$
$$\mathrm{E}\{a_2\}=\tfrac{1}{4}(8+7+12+23)=\$\mathbf{12\,500}\leftarrow \text{最优}$$
$$\mathrm{E}\{a_3\}=\tfrac{1}{4}(21+18+12+21)=\$18\,000$$
$$\mathrm{E}\{a_4\}=\tfrac{1}{4}(30+22+19+15)=\$21\,500$$

极小化极大准则 极小化极大准则产生下面的矩阵:

	s_1	s_2	s_3	s_4	行最大值
a_1	5	10	18	25	25
a_2	8	7	12	23	23
a_3	21	18	12	21	**21** ← 极小化极大值
a_4	30	22	19	15	30

损失后悔准则 从 1, 2, 3, 4 列中分别减去 5, 7, 12, 15, 得出后悔矩阵如下:

	s_1	s_2	s_3	s_4	行最大值
a_1	0	3	6	10	10
a_2	3	0	0	8	**8** ← 极小化极大值
a_3	16	11	0	6	16
a_4	25	15	7	0	25

赫尔维茨准则 下表给出计算结果:

行动方案	行最小值	行最大值	α(行最小值) + $(1-\alpha)$(行最大值)
a_1	5	25	$25-20\alpha$
a_2	7	23	$23-16\alpha$
a_3	12	21	$21-9\alpha$
a_4	15	30	$30-15\alpha$

用某个适当的 α, 就可以求出最优的行动方案. 例如当 $\alpha = 0.5$ 时, a_1 或 a_2 都产生最优决策; 而当 $\alpha = 0.25$ 时, a_3 最优.

Excel 程序

电子表格程序 excelUncertainty.xls 用来自动计算拉普拉斯、极大化极小、损失后悔和赫尔维茨准则. 该程序使用一个*费用*矩阵. 如要使用收益矩阵, 所有元素必须乘以 -1. 图 11.8 显示用本程序求解例 11.3-1, 最大矩阵维数为 10×10 (为了节省空间, 将列 F:K 隐藏).

	A	B	C	D	E	L	M	N	O
1	Decision Under Uncertainty								
2	Enter x to select method:					Output Results			
3		Laplace	x						
4		Minimax	x						
5		Savage	x						
6		Hurwicz	x	Alpha=	0.5	Optimum strategies			
7	Input (cost) Matrix: Maximum size = (10x10)					a2	a3	a2	a2
8		s1	s2	s3	s4	Laplace	Minimax	Savage	Hurwicz
9	a1	5	10	18	25	14.5	25	10	15
10	a2	8	7	12	23	12.5	23	8	15
11	a3	21	18	12	21	18	21	16	16.5
12	a4	30	22	19	15	21.5	30	25	22.5
13									
14									
15									
16									
17									
18									
19									

图 11.8　例 11.3-1 的 Excel 解 (文件 excelUncertainty.xls)

习题 11.3A

*1. Hank 是个聪明的学生, 每次考试前天晚上他只要温习一下课程往往都能得到好成绩. 但对明天的考试, Hank 遇到了一个小问题, 因为他想参加兄弟会哥们儿开的一个通宵派对. Hank 有 3 个选择:

$a_1 =$ 整晚上派对

$a_2 =$ 把这个晚上时间平均分两半, 一半学习, 一半派对

$a_3 =$ 整晚上都学习

明天的考试可能容易 (s_1)、适中 (s_2), 也可能很难 (s_3), 取决于教授难以揣测的情绪. Hank 期望得到下列分数:

	s_1	s_2	s_3
a_1	85	60	40
a_2	92	85	81
a_3	100	88	82

(a) (根据 4 个不确定型决策准则的每一个) 为 Hank 建议一个行动方案.

(b) 假设 Hank 对他将得到的用字母表示的分级成绩更感兴趣, 从 A 到 D 的成绩分别对应于 90, 80, 70, 60 分.(等级 F 表示不及格) 对这种成绩的态度会改变 Hank 的行动吗?

2. 对即将到来的播种季节, 农场主 McCoy 可以种玉米 (a_1), 也可以种小麦 (a_2), 或者种大豆 (a_3), 也可能把这块地用来放牧 (a_4). 相对不同行动的益损值受到降雨量的影响, 包括: 降雨量大 (s_1)、降雨量适中 (s_2)、降雨量小 (s_3) 或者干旱 (s_4).

估计的益损矩阵 (单位: 1 000 美元) 为

	s_1	s_2	s_3	s_4
a_1	−20	60	30	−5
a_2	40	50	35	0
a_3	−50	100	45	−10
a_4	12	15	15	10

给出农场主 McCoy 的行动方案.

3. 在 N 台机器中必须选出一台制造 Q 件某种产品. 对该种产品的最小和最大需求量分别是 Q^* 和 Q^{**}. 在机器 i 上生产 Q 件产品的生产费用包括一笔固定费用 K_i 和单位产品的可变费用 c_i 如下:

$$TC_i = K_i + c_iQ$$

(a) 应用 4 个不确定型决策准则的每一个, 求出本问题的一个解.

(b) 假设 $1\,000 \leqslant Q \leqslant 4\,000$, 用下列数据求解本问题:

机器 i	$K_i(\$)$	$C_i(\$)$
1	100	5
2	40	12
3	150	3
4	90	8

11.4 对 策 论

对策论所研究的决策情形中, 具有冲突目标的两个智能型对手都想要胜过对方. 典型的对策论例子包括竞争产品的广告战, 以及为作战部队制定战略等.

在对策冲突中, 两个对手称为**局中人** (player), 每个局中人都有若干 (有限或无限个) 方法, 或叫**策略**(strategy). 相应于每一对策略, 一个局中人从另一个局中人那里会得到一个**收益** (payoff). 这类对策被称为**二人零和对策** (two-person zero-sum game), 因为一个局中人得到的收益意味着另一个局中人付出等量的损失. 这时我们用一个局中人的收益来评判这种对策. 假定两个局中人 A 和 B 分别有 m 和 n 种策略, 那么这一对策通常可用局中人 A 的收益矩阵来表示:

	B_1	B_2	$\cdots$	B_n
A_1	a_{11}	a_{12}	$\cdots$	a_{1n}
A_2	a_{21}	a_{22}	$\cdots$	a_{2n}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
A_m	a_{m1}	a_{m2}	$\cdots$	a_{mn}

这种表示法说明, 如果 A 采用策略 i, 且 B 采用策略 j, 则 A 的收益为 a_{ij}, 同时意味着 B 的收益是 $-a_{ij}$.

实际应用——莱德杯比赛最后一天排定高尔夫球手比赛顺序

在高尔夫循环赛的最后一天, 两支球队争夺冠军. 每支球队的队长必须提交一份高尔夫球手的出场顺序单 (*出场名单*), 这个名单也就确定了比赛场次. 我们似乎可以假定, 假如两个水平相当的球手在他们的分别出场名单中排在同样的位置, 那么每个球手赢得比赛的机会为 50:50. 当排在前面的球手和排在后面的球手比赛时, 获胜的概率就会增加. 我们的目标是建立一套解析的方法, 看看采用比赛出场名单的想法是对还是不对. 第 24 章的案例 12 给出了这项研究的细节.

11.4.1　二人零和对策的最优解

由于对策来源于利益的冲突, 最优解就要每个局中人选择一个或多个策略, 使得对所选策略的任何改变都不改善双方局中人的收益. 这些解可以是单一纯粹策略的形式, 也可以是根据具体概率所得出的几个混合策略. 下面的例子说明了这两种情况.

例 11.4-1

两家公司 A 和 B 销售两个品牌的感冒药. 公司 A 的产品广告发布在广播 (A_1)、电视 (A_2) 和报纸 (A_3) 上, 公司 B 除了用广播 (B_1)、电视 (B_2) 和报纸 (B_3) 做广告外, 还寄发宣传页 (B_4). 根据每种广告宣传的效果, 每家公司都可以从另一方争夺一部分市场. 下面的矩阵给出了公司 A 所夺到的或丢失的市场份额百分比.

	B_1	B_2	B_3	B_4	行最小值
A_1	8	−2	9	−3	−3
A_2	6	**5**	6	8	**5** ← 极大化极小值
A_3	−2	4	−9	5	−9
列最大值	8	**5** ↑ 极小化极大值	9	8	

求解这一对策问题是基于保证每个局中人*最坏中的最好*原则. 假如公司 A 选

择策略 A_1, 则不论公司 B 采取什么策略, 可能发生的最坏情况就是 A 损失 3% 的市场份额给 B. 这可由第 1 行的最小值来表示. 类似地, 策略 A_2 的最坏结果是 A 从 B 那里夺到 5% 的市场, 策略 A_3 的最坏结果是 A 失去 9% 的市场给 B. 这些结果列在"行最小值"列下面. 为了得到*最坏中的最好结果*, 公司 A 选中了策略 A_2, 因为它对应于"行最小值"一列中的最大元素, 即极大化极小值.

接下来考虑公司 B 的策略. 因为给出的收益矩阵是对于公司 A 的, 公司 B 的*最坏中的最好*准则要求求出矩阵的极小化极大值. 其结果是, 公司 B 应该采取策略 B_2.

该对策的最优解要求选择策略 A_2 和 B_2, 也就是说, 这两家公司都应该利用电视做广告. 益损结果对公司 A 有利, 因为它的市场份额会增加 5%. 在这种情况下, 我们说**对策的价值** (value of the game) 是 5%, 并且 A 和 B 都在使用**鞍点** (saddle-point) 解.

鞍点解排除了任何一家公司对更好解的选择. 假如公司 B 换了另外一个策略 (B_1, B_3 或 B_4) 的话, 公司 A 可以仍然采用策略 A_2, 可以保证 B 将会失去更多的市场份额 (6% 或 8%). 同样的含义是, A 并不想要采用别的策略, 因为假如 A 换了策略 A_3, 则 B 可以换成 B_3, 导致增加市场份额 9%. 类似的结论是, 如果 A 采用 A_1, 则 B 可以采用 B_4 使得市场份额增加 3%.

一个对策的最优鞍点解并不一定是一个纯策略. 这个解可以是两个以上策略的随机混合, 如下面的例子所示.

例 11.4-2

有两个局中人 A 和 B 在玩掷硬币游戏. 每个局中人在不知道对方情况之下, 可选正面 (H) 或反面 (T). 这两个局中人同时公布他们的选择. 如果他们的结果相同 (HH 或 TT), 局中人 A 从局中人 B 得到 \$1. 否则 A 付给 B\$1.

下面 A 的益损矩阵给出了 2 行 2 列的值, 其行最小值和列最大值分别相应于 A 和 B 的策略.

	B_H	B_T	行最小值
A_H	1	−1	−1
A_T	−1	1	−1
列最大值	1	1	

该对策的极大化极小值和极小化极大值分别是 −\$1 和 \$1. 由于这两个值不相等, 因此该对策没有一个纯策略解. 特别是, 如果局中人 A 采用 A_H, 局中人 B 将会选择 B_T 而从 A 得到 \$1. 如果是这种情况, A 可以换成策略 A_T, 通过赢 \$1 使得对策结果反过来. 这种转向另外一种对策的不断诱惑表明, 一个纯策略解是不可接受的. 反之, 这两个局中人都可以随机把他们各自的纯策略混合起来. 在这种情况下, 这一对策的最优值会在对策的极大化极小值和极小化极大值之间的某个中间值取

得, 即

$$\text{极大化极小 (较低) 值} \leqslant \text{对策值} \leqslant \text{极小化极大 (较高) 值}$$

(见习题 11.4A 第 5 题) 因此, 在这个掷硬币例子中, 对策值一定在 −\$1 到 \$1 之间.

习题 11.4A

1. 求鞍点解、相应的纯策略以及下面每组对策的对策值. 益损值是对于局中人 A 的.

*(a)

	B_1	B_2	B_3	B_4
A_1	8	6	2	8
A_2	8	9	4	5
A_3	7	5	3	5

(b)

	B_1	B_2	B_3	B_4
A_1	4	−4	−5	6
A_2	−3	−4	−9	−2
A_3	6	7	−8	−9
A_4	7	3	−9	5

2. 以下给出 A 的益损值. 求使得每个对策的 (2,2) 元素成为一鞍点的 p 值和 q 值:

*(a)

	B_1	B_2	B_3
A_1	1	q	6
A_2	p	5	10
A_3	6	2	3

(b)

	B_1	B_2	B_3
A_1	2	4	5
A_2	10	7	q
A_3	4	p	6

3. 指出下列每种情况下对策值的区间, 假设益损值是对于局中人 A 的.

*(a)

	B_1	B_2	B_3	B_4
A_1	1	9	6	0
A_2	2	3	8	4
A_3	−5	−2	10	−3
A_4	7	4	−2	−5

(b)

	B_1	B_2	B_3	B_4
A_1	−1	9	6	8
A_2	−2	10	4	6
A_3	5	3	0	7
A_4	7	−2	8	4

(c)

	B_1	B_2	B_3
A_1	3	6	1
A_2	5	2	3
A_3	4	2	−5

(d)

	B_1	B_2	B_3	B_4
A_1	3	7	1	3
A_2	4	8	0	−6
A_3	6	−9	−2	4

4. 两家公司推出两种竞争产品. 目前, 每种产品占据 50% 的市场. 由于最近对这两种产品的改进, 每家公司都打算开展广告促销. 假如两家公司都不做广告的话, 将继续平分市场. 但如果任何一家公司发布更强势的广告宣传, 另一家公司当然就会损失其一部分的客户. 一项市场调查发现, 通过电视广告可以宣传到 50% 的潜在顾客, 通过报纸可宣传到 30% 的潜在顾客, 而通过广播则可以宣传到 20% 的潜在顾客.

(a) 将该问题表达为一个二人零和对策, 并为每家公司选择合适的广告媒体.

(b) 求出该对策值的区间. 每家公司能按照某种单一纯策略进行广告宣传吗?

5. 设局中人 A 有 m 个策略, 局中人 B 有 n 个策略. A 的益损矩阵的第 (i, j) 元素为 a_{ij}. 证明

$$\max_i \min_j a_{ij} \leqslant \min_j \max_i a_{ij}$$

11.4.2 求解混合策略对策

带有混合策略的对策可以用作图法求解, 也可以用线性规划求解. 图解法适用于至少有一个局中人恰好有两个纯策略的对策问题. 图解法非常有意思, 因为它用图形解释了鞍点的含义, 而线性规划可以用来求解任何一个二人零和对策.

对策的图解法 从 $(2\times n)$ 对策的情形开始, 其中局中人 A 有两个策略.

		y_1	y_2	$\cdots$	y_n
		B_1	B_2	$\cdots$	B_n
x_1 :	A_1	a_{11}	a_{12}	$\cdots$	a_{1n}
$1-x_1$:	A_2	a_{21}	a_{22}	$\cdots$	a_{2n}

这个对策中, 局中人 A 分别用概率 x_1 和 $1-x_1, 0\leqslant x_1\leqslant 1$ 混合了策略 A_1 和 A_2, 局中人 B 用概率 $y_1, y_2, \cdots$ 和 y_n 混合了策略 B_1 到 B_n, 其中 $y_j\geqslant 0, j=1,2,\cdots,n$, 并且 $y_1+y_2+\cdots+y_n=1$. 在这种情况下, 对应于 B 的第 j 个纯策略, A 的期望收益可计算为

$$(a_{1j}-a_{2j})x_1+a_{2j},\quad j=1,2,\cdots,n$$

因此局中人 A 只要求出 x_1 的值, 使得最小的期望收益达到最大, 即

$$\max_{x_1}\min_{j}\{(a_{1j}-a_{2j})x_1+a_{2j}\}$$

例 11.4-3

考虑以下的 2×4 对策, 局中人 A 的收益为

	B_1	B_2	B_3	B_4
A_1	2	2	3	−1
A_2	4	3	2	6

这个对策没有纯策略解. A 对应于 B 的纯策略的期望收益为

B 的纯策略	A 的期望收益
1	$-2x_1+4$
2	$-x_1+3$
3	x_1+2
4	$-7x_1+6$

图 11.9 给出了 TORA 画出的对应于 B 的纯策略的 4 条直线 (文件 toraEx11.4-3.txt)[①]. 为了确定出**最坏中的最好解**, 4 条直线的下包络线 (垂直竖条表示) 代表了在 B 采取任一策略时 A 的最小 (最坏) 的期望收益, 而下包络线的最大值 (最好

① 从 Main menu 下选择 Zero-sum Games 并输入问题数据, 然后在 SOLVE/MODIFY 菜单中选择 Graphical .

值) 对应于在 $x^* = 0.5$ 的极大化极小值解点, 这个点是相对于策略 B_3 和 B_4 直线的交点. 局中人 A 的最优解要求混合策略 A_1 和 A_2, 概率分别为 0.5 和 0.5.

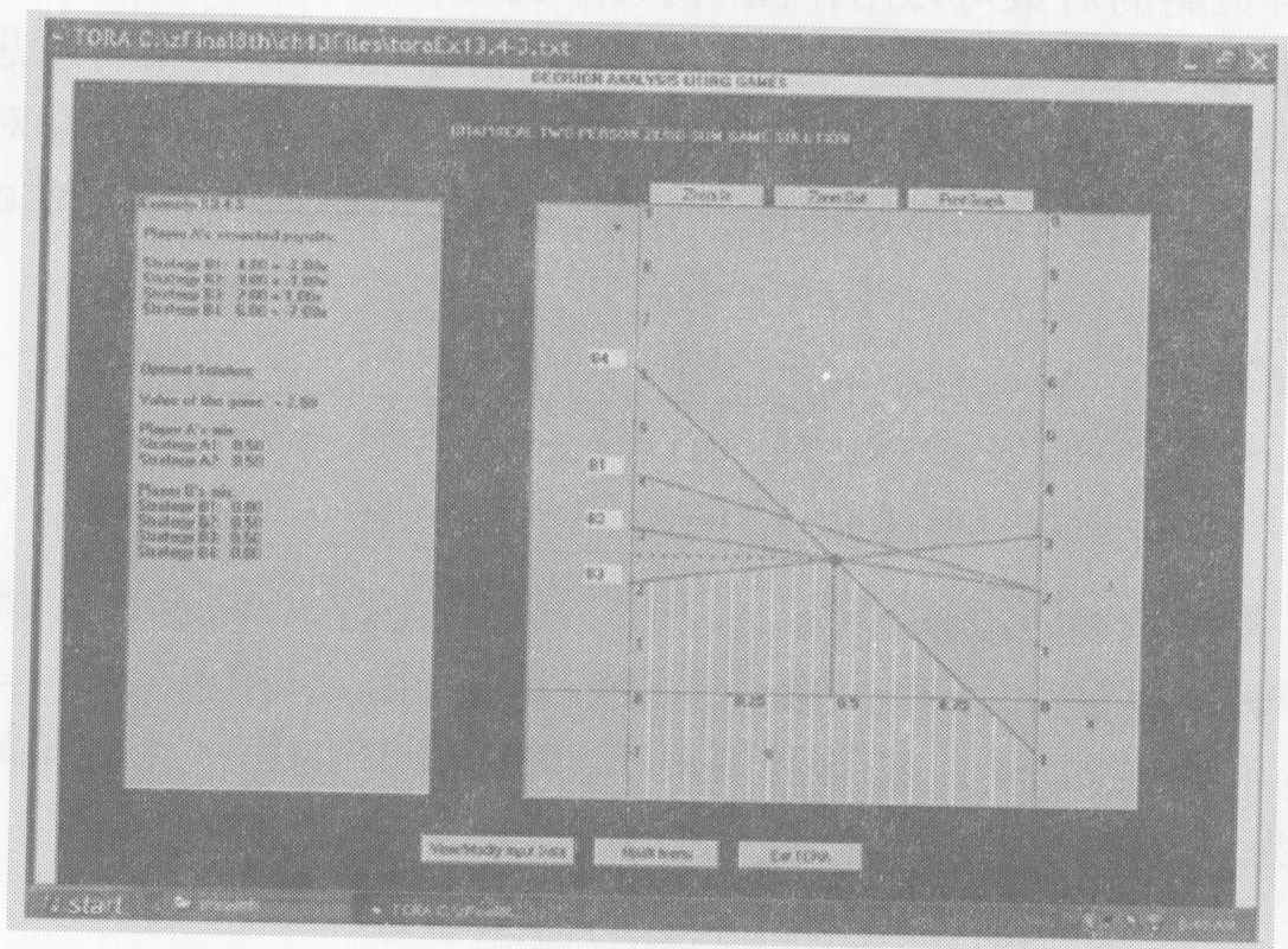


图 11.9　例 11.4-3 二人零和对策的 TORA 图解 (文件 toraEx11.4-3.txt)

求相应的对策值 v, 只要从直线 3 或直线 4 的方程中减去 $x_1 = 0.5$, 我们得到

$$v = \begin{cases} \frac{1}{2} + 2 = \frac{5}{2}, & \text{从直线 3} \\ -7\left(\frac{1}{2}\right) + 6 = \frac{5}{2}, & \text{从直线 4} \end{cases}$$

由图中下包络线定义的两个策略可求出局中人 B 的混合策略. 这意味着 B 可以混合策略 B_3 和 B_4, 在这种情况下 $y_1 = y_2 = 0$, 且 $y_4 = 1 - y_3$. 结果是, 对应 A 的纯策略的 B 的期望收益为

A 的纯策略	B 的期望收益
1	$4y_3 - 1$
2	$-4y_3 + 6$

B 的**最坏中的最好**解为所给出的两条直线的上包络线的最小点 (你会发现, 画出两条直线并找出上包络线是非常有用的). 这个过程相当于求解方程

$$4y_3 - 1 = -4y_3 + 6$$

方程的解为 $y_3 = \frac{7}{8}$, 得出的对策值为 $v = 4 \times \left(\frac{7}{8}\right) - 1 = \frac{5}{2}$.

这个对策的解要求局中人 A 等概率混合策略 A_1 和 A_2, 局中人 B 以概率 $\frac{7}{8}$ 和 $\frac{1}{8}$ 混合策略 B_3 和 B_4. (实际上, 对策中局中人 B 还有另一个解, 因为图 11.9 中的极大化极小值点可由多于两条直线确定, 这些不同解的任何非负组合也是一个可行解.)

评注　可类似地处理局中人 A 有 m 个策略、局中人 B 只有两个策略的对策问题. 主要的区别是, 我们画出对应于 A 的纯策略的 B 的期望收益. 这样, 我们需要求出

所画直线的上包络线的极小化极大值点而不是极大化极小值点. 但为了用 TORA 求解, 必须要用 -1 乘以收益矩阵来表示有两个策略的局中人的收益.

习题 11.4B[①]

***1.** 用图解法求解例 11.4-2 的掷硬币对策.

***2.** Robin 经常往返于两个城市之间, 他有两条路线可选择: 路线 A 是一条 4 车道的高速路, 路线 B 是一条长弯路. 高速公路巡警的警力有限. 假如全部警力都布置在一条路线上, 因为 Robin 热衷于开"超快"车, 他肯定会受到超速罚款 \$100. 如果巡警分成两半各布置在两条路上, 则他在路线 A 上受到罚款的机率为 50%, 而在路线 B 上受到同样数额罚款的机率为 30%. 为 Robin 和公路巡警都提出一个策略.

3. 用图解法解下列对策, 益损值是针对局中人 A 的.

(a)

	B_1	B_2	B_3
A_1	1	-3	7
A_2	2	4	-6

(b)

	B_1	B_2
A_1	5	8
A_2	6	5
A_3	5	7

4. 考虑下面的二人零和对策:

	B_1	B_2	B_3
A_1	5	50	50
A_2	1	1	0.1
A_3	10	1	10

(a) 证明 A 的最优策略是 $(\frac{1}{6}, 0, \frac{5}{6})$, B 的最优策略为 $(\frac{49}{54}, \frac{5}{54}, 0)$, 并求出对策值.

(b) 说明最优对策值等于

$$\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}x_i y_j$$

对策的线性规划求解 对策论与线性规划有着密切的联系, 因为一个二人零和对策可以表示成一个线性规划, 反之亦然. 事实上, G. Dantzig (1963, p.24) 里写道, 在 1947 年首次引入单纯形方法之初, 对策论之父冯 · 诺伊曼就立即发现了它们之间的这种关系, 并进一步提出和强调了线性规划中的对偶概念. 本节介绍如何用线性规划求解对策问题.

局中人 A 的最优对策 $x_1, x_2, \cdots, x_m$ 可以从下面的求极大化极小值问题得到:

$$\max_{x_i}\left\{\min\left(\sum_{i=1}^{m} a_{i1}x_i, \sum_{i=1}^{m} a_{i2}x_i, \cdots, \sum_{i=1}^{m} a_{in}x_i\right)\right\}$$

$$\text{s.t.}\quad x_1 + x_2 + \cdots + x_m = 1$$

$$x_i \geqslant 0,\quad i = 1, 2, \cdots, m$$

① 可用 TORA 的 Zero-sum Games 程序模块来检查你的答案.

现在令

$$v=\min\left\{\sum_{i=1}^{m}a_{i1}x_i,\sum_{i=1}^{m}a_{i2}x_i,\cdots,\sum_{i=1}^{m}a_{in}x_i\right\}$$

从这一方程可得到

$$\sum_{i=1}^{m}a_{ij}x_i\geqslant v,\quad j=1,2,\cdots,n$$

因此局中人 A 的问题可写成

$$\begin{aligned}\max\quad & z=v\\ \text{s.t.}\quad & v-\sum_{i=1}^{m}a_{ij}x_i\leqslant 0,\quad j=1,2,\cdots,n\\ & x_1+x_2+\cdots+x_m=1\\ & x_i\geqslant 0, i=1,2,\cdots,m\\ & v\text{无限制}\end{aligned}$$

注意到对策值 v 的符号没有限制.

局中人 B 的最优策略 $y_1,y_2,\cdots,y_n$ 可以从解下面的问题得到:

$$\begin{aligned}&\min_{y_j}\left\{\max\left(\sum_{j=1}^{n}a_{1j}y_j,\sum_{j=1}^{n}a_{2j}y_j,\cdots,\sum_{j=1}^{n}a_{mj}y_j\right)\right\}\\ &\text{s.t. } y_1+y_2+\cdots+y_n=1\\ &\qquad y_j\geqslant 0,\quad j=1,2,\cdots,n\end{aligned}$$

运用与局中人 A 类似的方法, 局中人 B 的问题可简化为

$$\begin{aligned}\min\quad & w=v\\ \text{s.t.}\quad & v-\sum_{j=1}^{n}a_{ij}y_j\geqslant 0,\quad i=1,2,\cdots,m\\ & y_1+y_2+\cdots+y_n=1\\ & y_j\geqslant 0,\quad j=1,2,\cdots,n\\ & v\text{无限制}\end{aligned}$$

这两个问题对同样的 (无限制) 对策值变量 v 求最优, 其原因是, B 的问题是 A 的问题的对偶 (用第 4 章 对偶性的定义证明). 这就意味着, 一个问题的最优解自动地成为了另一个问题的最优解.

例 11.4-4

用线性规划求解下列对策问题.

	B_1	B_2	B_3	行最小值
A_1	3	−1	−3	−3
A_2	−2	4	−1	**−2**
A_3	−5	−6	2	−6
列最大值	3	4	**2**	

对策值 v 在 -2 到 2 之间.

局中人 A 的线性规划问题

$$\begin{aligned}\max \quad & z = v \\ \text{s.t.} \quad & v - 3x_1 + 2x_2 + 5x_3 \leqslant 0 \\ & v + x_1 - 4x_2 + 6x_3 \leqslant 0 \\ & v + 3x_1 + x_2 - 2x_3 \leqslant 0 \\ & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geqslant 0 \\ & v\text{无符号限制}\end{aligned}$$

最优解① 为 $x_1 = 0.39, x_2 = 0.31, x_3 = 0.29, v = -0.91$.

局中人 B 的线性规划问题

$$\begin{aligned}\min \quad & z = v \\ \text{s.t.} \quad & v - 3y_1 + y_2 + 3y_3 \geqslant 0 \\ & v + 2y_1 - 4y_2 + y_3 \geqslant 0 \\ & v + 5y_1 + 6y_2 - 2y_3 \geqslant 0 \\ & y_1 + y_2 + y_3 = 1 \\ & v\text{无符号限制}\end{aligned}$$

最优解为 $y_1 = 0.32, y_2 = 0.08, y_3 = 0.60, v = -0.91$.

习题 11.4C

1. 在一次野餐会上, 2 个 2 人组在玩捉迷藏游戏. 共有 4 个隐藏地点 (A, B, C, D), 隐藏组的 2 名成员可以分别藏在 4 个地点的任何 2 个, 另一组人有机会寻找任何 2 个地点. 如果他们都找到了隐藏组的 2 个人, 搜寻组就可以得到一分奖励; 假如没有把两个人都找到, 他们就输一分. 其他情况下, 结果是平局.

*(a) 把这个问题表示成一个二人零和对策. (b) 求出最优策略和对策值.

2. UA 和 DU 在为 2008 年全国大学篮球锦标赛制定他们的策略. 在评估了他们分别的“板凳队员”的强势后, 每队的教练都提出了比赛中替补队员的 4 个策略. 每队 2 分球手、3 分球手和罚球手的能力如何就成为决定最终比赛成绩的关键因素. 下表给出了 UA 作为每队不同策略函数的每次发球将会得到的净分数:

① TORA Zero-sum Games ⇒ Solve ⇒ LP-based 可用来求解任何二人零和对策.

	DU_1	DU_2	DU_3	DU_4
UA_1	3	−2	1	2
UA_2	2	3	−3	0
UA_3	−1	2	−2	2
UA_4	−1	−2	4	1

(a) 用线性规划求解该对策问题, 并确定锦标赛的一个策略.

(b) 根据已有的信息, 这两个队哪个可能赢得锦标赛?

(c) 假设整个比赛有 60 次发球 (每队 30 次), 预测锦标赛获胜队的期望分数.

3. Blotto 上校的部队在为控制 2 个战略要地而战斗. Blotto 有 2 个团, 而敌人有 3 个团. 投入军团多的部队将会占领 1 个要地, 否则势均力敌.

*(a) 将该问题表示成一个二人零和对策并用线性规划求解.

(b) 哪支部队将会赢得这场战斗?

4. 在 2 人用 2 个手指猜拳游戏中, 每个做游戏的人出 1 个或 2 个手指, 并同时猜对手出示的手指数. 猜对了的赢得与双方所出手指总数相等的分数, 否则游戏平局. 将这个问题表示成一个二人零和对策, 并用线性规划求解.

参 考 文 献

Chen, S., and Hwang, C., *Fuzzy Multiple Attribute Decision Making*, Springer-Verlag, Berlin, 1992.

Clemen, R.J., and Reilly, T., *Making Hard Decisions: An Introduction to Decision Analysis*, 2nd ed., Duxbury, Pacific Grove, CA, 1996.

Cohan, D., S, Haas, D. Radloff, and R. Yancik, "Using Fire in Forest Management: Decision Making under Uncertainty," *Interfaces*, Vol. 14, No. 5, pp. 8-19, 1984.

Dantzig, G.B., *Linear Programming and Extensions*, Princeton University Press, Princeton, NJ, 1963.

Meyerson, R., *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA, 1991.

Rapport, A., "Sensitivity Analysis in Decision Making," *The Accounting Review*, Vol. 42, No. 3, pp. 441-456, 1967.

Saaty, T.L., *Fundamentals of Decision Making*, RWS Publications, Pittsburgh, 1994.

Zahedi, F., "The Analytic Hierarchy Process - A Survey of the Method and its Applications," *Interfaces*, Vol. 16, No. 4, pp. 96-108, 1986.

Zeleny, M., *Multiple Criteria Decision Making*, McGraw-Hill, New York, 1982.

第12章 排队系统

本章导读 分析排队系统的目的是为了对排队等待的顾客提供满意的服务. 和前面各章介绍的 OR 工具不同, 排队论并不是一种优化技术, 而是用于度量排队系统的性能指标, 如队列的平均等待时间和服务设施的效率, 这些度量指标可以用来设置服务设施. 本章重点在于实际中排队分析结果的实施. 但为了充分理解排队系统的实际问题, 就需要了解相当的基础理论背景. 为此, 本章先介绍两个重要分布 (泊松分布和指数分布) 的"完全随机"性质. 这一点是非常重要的, 因为这会帮助我们了解实际中排队结果应用的具体情形.

排队结果涉及计算难度很大的公式, 建议你用 excelPoissonQ.xls 或 TORA 程序来做这些计算. 你会发现 TORA 在比较多种情景方面非常有用. 本章自始至终都用 TORA 进行计算, 并对这些结果的实际解释进行大量的讨论. 建议你在求解本章习题时也采用同样的方法, 这样你就不会陷于繁琐的计算细节, 并能方便地对不同的情景进行试验.

本章包括 2 个实际应用、17 个例子、2 个 Excel 程序、137 个节后习题以及 5 个案例. 这些案例收集在附录 E 中. AMPL/Excel/Solver/TORA 程序放在文件夹 ch12Files 下.

实际应用——制造厂内部运输系统的分析

某制造厂用 3 辆卡车运输材料. 这些卡车都在中心停车场等待发车请求. 一辆收到发车请求的卡车要开到客户地点, 装车后开到目的地, 然后再返回到中心停车场. 这项服务的主要用户包括生产部门、制造车间和维修部门. 其他部门偶尔也会请求使用这些卡车.

用户因为等待空车时间过长而提出了抱怨, 特别是生产部门, 他们要求车队增加第 4 辆卡车. 但这项申请并不合理, 因为排队论分析表明, 等待时间过长主要是物流造成的, 只要改变一下车队运行的规程, 并不需要增加第 4 辆卡车. 第 24 章案例 14 给出了这一研究的细节.

12.1 为什么要研究排队系统

日常生活中常常需要等待服务. 例如在餐馆等候服务, 在商店交费柜台前"排队等候", 在邮局"排队"等待服务. 但是排队现象不仅仅是我们人类独有的, 比如

工件等待机器加工, 飞机在机场上空盘旋等待批准着陆, 汽车等待交通信号等, 它们也存在着排队现象. 不花费极大的成本, 等待现象是不可能完全消除的, 我们的目标是要把它的不利影响减小到"可以忍受的"程度.

排队论主要运用像平均队列长度、平均等待时间, 以及设施平均利用率这样的性能度量指标, 来定量研究排队现象. 下例说明了如何运用这些度量来设计服务设施.

例 12.1-1

McBurger 是一家快餐店, 有 3 个服务柜台. 这家店的经理委托他人调查顾客对服务速度慢的投诉. 调查研究结果显示, 服务柜台数量与服务等待时间之间有着如下关系:

收款台数	1	2	3	4	5	6	7
平均等待时间 (分钟)	16.2	10.3	6.9	4.8	2.9	1.9	1.3

仔细观察这些数据发现, 在现有 3 个柜台的情况下, 平均等待时间要 7 分钟. 需要 5 个柜台才能把等待时间减少到 3 分钟.

评注 排队分析的结果可以用在费用优化模型中, 即求两种费用 (提供服务的费用和等待的费用) 之和的最小值. 图 12.1 显示了一个典型的费用模型 (每单位时间费用), 其中服务费用随着服务水平 (例如, 服务柜台数) 的提高而增加, 同时, 等待时间随着服务水平的提高而减少. 使用费用模型的主要障碍是很难估计可靠的等待费用, 特别当人的行为成为操作的有机组成部分时. 我们在 12.9 节讨论这一问题.

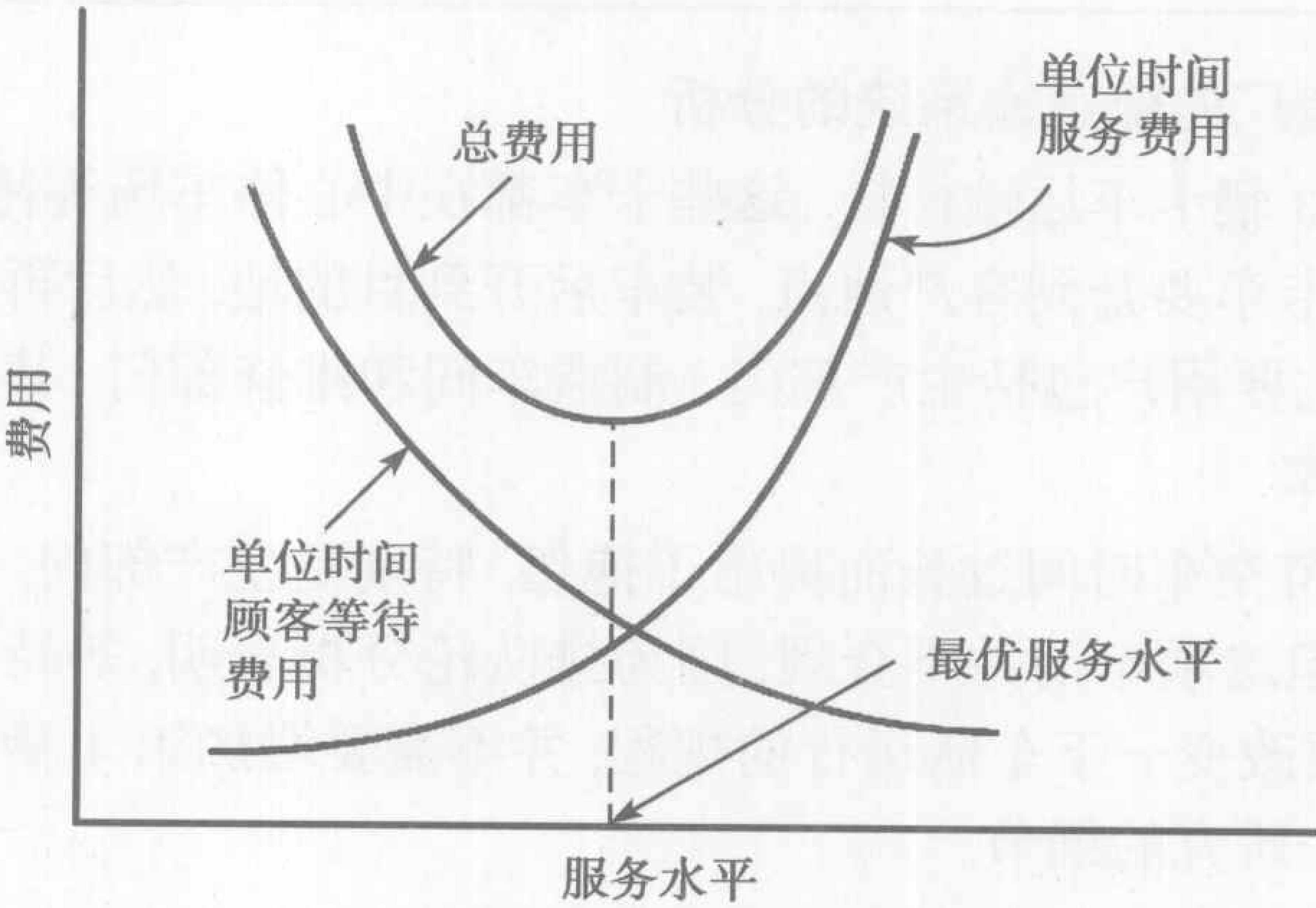


图 12.1 基于费用的排队决策模型

习题 12.1A

*1. 假定对 McBurger 快餐店的进一步分析显示以下结果:

收款台数	1	2	3	4	5	6	7
空闲率 (%)	0	8	12	18	29	36	42

(a) 当收款柜台数为 5 个时, 该快餐店的生产率是多少 (用员工繁忙时间的百分比来表示)?

(b) 经理想要把平均等待时间保持在 3 分钟左右, 同时维持服务设施的效率约达到 90%. 这两个目标能都达到吗? 请解释.

2. Acme 金属工件加工厂正在采购一种多用途钻床, 有两个型号 A 和 B 可选, 每小时加工成本分别为 $18 和 $25. A 型钻床比 B 型钻床慢. 对类似机器的排队分析表明, 假如用 A 型钻床, 等待加工工件的平均数是 4 个, 比 B 型钻床的等待工件数多 30%. 而工件的延迟就代表着收入的损失. 根据 Acme 的估算, 每小时每个等待工件的费用为 $10, Acme 该买哪种型号的钻床?

12.2 排队模型的要素

一个排队系统中的主要参与者是**顾客** (customer) 和**服务台** (server). 顾客从某个**输入源** (source) 产生, 到达一个服务**设施** (facility), 他们可以立即得到服务; 假如服务设施繁忙, 也可能在**队列** (queue) 中等待. 当一个设施完成一次服务, 如果有顾客等待的话, 则自动地“拉出”一个等待顾客; 假如队列为空, 设施就变成空闲, 直到一个新的顾客到达.

从分析队列的角度, 我们用连续两个顾客之间的**到达间隔时间** (interarrival time) 表示顾客的到达, 用对每个顾客的**服务时间** (service time) 来描述服务. 一般地, 到达间隔时间和服务时间可以随机的, 如邮局的服务系统; 也可以是确定性的, 如求职面试申请者的到达.

队列长度 (queue size) 对于队列的分析有作用, 它可以是有限长的, 如两个相邻机器之间的缓冲区; 也可以是无限的, 如邮寄订单处理.

排队规则 (queue discipline) 表示从队列里选择顾客的顺序, 是排队模型分析的一个重要因素. 最常见的排队规则是**先到先服务** (first come, first served, FCFS). 其他的排队规则还有**后到先服务** (last come, first served, LCFS) 和**随机顺序服务** (service in random order, SIRO). 也可以按照某种**优先权** (priority) 顺序从队列里挑选顾客, 例如车间里把紧急工件放在普通工件前面进行处理.

在排队分析中, 顾客的排队行为也起着作用. “人类”顾客可能会从一个队列**跳到**另一个队列, 以期望缩短等待时间. 顾客也可能会由于预计的排队时间过长而**暂时不加入**队列, 或者可能会从一个队列中等待过久而**退出**, 因为已经等待了太长的时间.

服务设施的设计可以包括并行服务 (如邮局或银行服务). 服务人员也可以安排成串行的 (如几台机器对工件的连续加工作业), 或者可以被连成网络 (如计算机

网络中的路径).

产生顾客的输入源可以是有限的, 也可以是无限的. **有限输入源** (finite source) 限制等候服务的顾客 (如要求修理人员服务的机器), 而**无限输入源** (infinite source) 则始终是充分多的 (如到达电话交换机的呼叫).

一个排队系统要素的各种变化导致了各种各样的排队模型. 本章会给出这些模型的例子.

习题 12.2A

1. 在下列每种情形中, 找出顾客和服务台.

*(a) 飞机抵达一个机场.　　*(b) 出租车站为等待的旅客提供服务.

(c) 从某加工车间的仓库里领出工具.　　(d) 邮局处理信件.

(e) 大学注册班级.　　(f) 法院审理案件.

(g) 超市付款.　　*(h) 停车场的运行.

2. 对第 1 题的每种情形, 找出以下内容: (a) 顾客源的性质 (有限或无限); (b) 到达的顾客的性质 (单个到达还是成批到达); (c) 到达间隔时间的类型 (随机性的还是确定性的); (d) 服务时间的定义和类型; (e) 队列容量 (有限队列还是无限队列); (f) 排队规则.

3. 研究下列系统并找出相应的排队情形. 对于每种情形, 定义顾客、服务台、排队规则、服务时间、最大队列长度以及顾客源.

某车间收到工件加工订单后, 负责人决定该订单是紧急工作还是一般工作. 某些订单只要求用一台机床, 而剩下的订单都要使用两台机床, 按照两级生产线进行加工. 每一组里, 专门有一套设备来处理紧急工作.

到达机床的工件按照其到达顺序进行加工, 完成加工的订货到达货场后就运走, 货场容量有限.

用于不同机床的打磨好的工具由中心工具库提供. 如果机床出现故障, 就从服务部门叫来修理工进行修理. 加工紧急订单的机床总是优先从工具库领到新工具, 并且优先得到维修服务.

4. 判断正误:

(a) 一个耐心不足的等待顾客可能决定退出队列.

(b) 假如估计的等待时间很长, 一个来到的顾客可能暂时不加入排队.

(c) 从一个队列跳到另一个队列一定能够减少等待时间.

5. 在第 1 题的每个情形下, 讨论顾客跳队、暂时不加入和退出排队的可能性.

12.3　指数分布的作用

在大多数排队情形中, 顾客的到达*完全是随机*发生的. 这里的随机性意味着, 一个事件的发生 (如一个顾客的到达或一项服务的完成) 不受上一个事件发生以后所经过的时间长度的影响.

在排队模型中，随机到达间隔时间和服务时间用指数分布来进行定量描述，定义为

$$f(t)=\lambda\mathrm{e}^{-\lambda t},\quad t>0$$

14.4.3 节表明，对于**指数分布** (exponential distribution)，

$$\mathrm{E}\{t\}=\frac{1}{\lambda}$$

$$\begin{aligned}\mathrm{P}\{t\leqslant T\}&=\int_0^T\lambda\mathrm{e}^{-\lambda t}\mathrm{d}t\\&=1-\mathrm{e}^{-\lambda T}\end{aligned}$$

$\mathrm{E}\{t\}$ 的定义表明，λ 是每单位时间产生事件 (到达或离去) 的速率. 我们用下面的例子来说明指数分布是**完全随机** 的这一事实：假定现在的时间是上午 8:20，上一个顾客的到达时间是 8:02，下一个到达发生在 8:29 之前的概率只是 8:20~8:29 这一区间的函数，与上一个事件的发生 (8:02~8:20) 以来所流逝的时间长度完全无关. 这一结果我们称为指数分布的**遗忘性**(forgetfulness) 或者叫**无记忆性** (lack of memory).

令指数分布 $f(t)$ 表示相继事件之间的时间 t 的分布. 如果 S 为上一个事件发生以来的时间区间，则遗忘性意味着

$$\mathrm{P}\{t>T+S|t>S\}=\mathrm{P}\{t>T\}$$

为了证明这个结果，注意到对于平均值为 $\frac{1}{\lambda}$ 的指数函数，我们有

$$\mathrm{P}\{t>Y\}=1-\mathrm{P}\{t<Y\}=\mathrm{e}^{-\lambda Y}$$

因此，

$$\begin{aligned}\mathrm{P}\{t>T+S|t>S\}&=\frac{\mathrm{P}\{t>T+S,t>S\}}{\mathrm{P}\{t>S\}}=\frac{\mathrm{P}\{t>T+S\}}{\mathrm{P}\{t>S\}}\\&=\frac{\mathrm{e}^{-\lambda(T+S)}}{\mathrm{e}^{-\lambda S}}=\mathrm{e}^{-\lambda T}\\&=\mathrm{P}\{t>T\}\end{aligned}$$

例 12.3-1

一台服务机器任何时候发生故障时，总有一台备用机立即替换. 该机器 (以及备份机) 发生故障的时间服从指数分布，并平均 5 小时发生一次. 机器操作员称，这机器“习惯于”每天晚上 8:30 坏掉，请分析该操作员的话.

该机器的平均故障率为 $\lambda=\frac{1}{5}=0.2$ 次故障/小时，因此，故障时间的指数分布为

$$f(t)=0.2\mathrm{e}^{-0.2t},\quad t>0$$

关于操作员的话, 我们马上知道是不对的, 因为它和故障之间的时间服从指数分布因此是完全随机的这一事实是矛盾的. 故障发生在晚上 8:30 之前的概率不能用来支持或驳斥操作员声称的结果, 因为这样的概率值依赖于当天 (相对于 8:30 P.M.) 计算概率值的时间. 例如, 若现在的时间是 8:20 P.M., 操作员今晚的声称是对的概率为

$$\mathrm{P}\left\{t<\frac{10}{60}\right\}=1-\mathrm{e}^{-0.2\left(\frac{10}{60}\right)}=0.032\ 78$$

这个值过低. 如果现在的时刻是 1:00 P.M., 则 8:30 P.M. 之前发生故障的概率增加到约 0.777(请证明!). 这两个极端值说明操作员的话是站不住脚的.

习题 12.3A

1. **(a)** 请解释你对到达率 λ 与平均到达间隔时间之间关系的理解. 描述每个变量的单位是什么?

(b) 对于下列每种情况, 求每小时平均到达率 λ 和平均到达间隔时间小时数.

*(i) 每 10 分钟发生一次到达. (ii) 每 6 分钟发生两次到达.

(iii) 30 分钟内的到达数为 10. (iv) 相继到达之间的平均间隔时间为 0.5 小时.

(c) 对于下列每种情况, 求每小时平均服务率 μ 和平均服务时间小时数.

*(i) 每 12 分钟完成一次服务. (ii) 每 15 分钟有两人离开.

(iii) 30 分钟内服务的顾客数为 5. (iv) 平均服务时间为 0.3 小时.

2. 在例 12.3-1 中, 求下列各值:

(a) 1 周内的平均故障数, 假定每周 7 天, 每天 24 小时提供服务.

(b) 2 小时内至少发生一次故障的概率.

(c) 3 小时内不发生下一次故障的概率.

(d) 若上一次故障后 3 小时内没有发生故障, 则故障间隔时间至少为 4 个小时的概率是多少?

3. 国家财政局 (State Revenue Office) 的到达间隔时间服从平均值 0.05 小时的指数分布. 该局上午 8:00 开始办公.

***(a)** 写出描述该到达间隔时间的指数分布.

***(b)** 求出到 8:15 A.M. 之前该局无顾客到达的概率.

(c) 现在是 8:35 A.M., 上一个顾客 8:26 进入该局, 下一个顾客在 8:38 A.M. 之前到达的概率是多少? 下一个顾客 8:40 A.M. 前仍未到达的概率是多少?

(d) 8:10~8:45 A.M 到达顾客的平均数是多少?

4. 假设一台机器故障间隔时间服从平均值为 6 小时的指数分布. 若该机器在前 3 个小时内无故障正常工作, 那么下一个小时内该机器仍连续无故障的概率是多少? 未来 0.5 小时内发生故障的概率是多少?

5. 学生会游戏室的到达间隔时间服从指数分布, 平均值为 10 分钟.

(a) 每小时到达率是多少?

(b) 求下 15 分钟内没有学生来到游戏室的概率.

(c) 下 20 分钟内至少有一个学生来到游戏室的概率是多少?

6. 一家新开张快餐店的经理想要定量分析以下顾客的到达过程, 他估计了几个顾客到达间隔时间: (a) 小于 2 分钟; (b)2 分钟到 3 分钟; (c) 多于 3 分钟. 同类快餐店的顾客到达率为每小时 35 名顾客. 到达间隔时间服从指数分布.

***7.** Ann 和 Jim 是一家快餐店的两名员工, 在等待顾客上门的时候他俩在玩一个游戏: 假如在 1 分钟内下一个顾客没来, 则 Jim 付给 Ann 2 分钱; 否则的话, Ann 付给 Jim 2 分钱. 求 8 小时内 Jim 的平均收益是多少. 到达间隔时间服从平均值 1.5 分钟的指数分布.

8. 假设在第 7 题中, 游戏规则改为: 假如在 1.5 分钟以后下一个顾客到达, 则 Jim 付给 Ann 2 分钱. 假如在 1 分钟内下一个顾客到达, 则 Ann 给 Jim 2 分钱. 如顾客在 1 分钟到 1.5 分钟之间到达, 则不分输赢. 求 8 小时内 Jim 的平均收益是多少.

9. 假设在第 7 题中, 假如下一个顾客在一分钟内到达, 则 Ann 给 Jim 2 分钱. 如相继到达间隔时间在 1 分钟到 1.5 分钟, 则 Ann 付给 Jim 3 分钱. 假如相继到达间隔时间在 1.5 分钟至 2 分钟之间, 则 Jim 付给 Ann 5 分钱. 如相继到达间隔时间大于 2 分钟, 则 Jim 付给 Ann 6 分钱. 求 8 小时内 Ann 的平均收益是多少.

***10.** 在前一名顾客后 4 分钟内来到 McBurger 快餐店的顾客将会得到 10% 的优惠. 如果间隔时间是 4 到 5 分钟, 优惠幅度为 6%. 假如间隔时间在 5 分钟以上, 则该顾客只能得到 2% 的优惠. 间隔时间是服从均值为 6 分钟的指数分布.

(a) 求顾客得到 10% 优惠的概率是多少? **(b)** 求顾客的平均优惠幅度是多少?

11. 已知一台 Kencore 冰箱的故障间隔时间服从指数分布, 平均值为 9 000 小时 (大约 1 年的运行时间). 公司为冰箱提供 1 年的保修期, 那么由保修承担故障维修费的机会有多少呢?

12. A 大学运行两条校内巴士线路: 红线和绿线. 红线服务北校区, 绿线服务南校区, 并由一个换乘站相连. 绿线巴士每 10 分钟随机 (指数到达间隔时间) 到达换乘站, 红线巴士每 7 分钟也随机到达.

(a) 一学生乘红线到达换乘绿线, 他等待时间的概率分布是什么?

(b) 一学生乘绿线到达换乘红线, 他等待时间的概率分布是什么?

13. 证明指数分布的平均值和标准差相等.

12.4 纯生模型和纯灭模型(指数分布和泊松分布之间的关系)

本节介绍两个排队情形: 只有到达的纯生 (pure birth) 模型和只发生离开的纯灭 (pure death) 模型. 纯生模型的例子如为新生婴儿制作出生证明, 纯灭模型的例子如一家商店对其库存货物的随机提货.

我们用指数分布来描述纯生模型中的到达间隔时间和纯灭模型中的离开间隔时间. 建立这两个模型, 结果却发现了指数分布与泊松分布之间的密切关系, 即由其中一个分布可以自动确定另一个分布.

12.4.1　纯生模型

定义

$$p_0(t) = t \text{ 时间期间内没有到达的概率}$$

已知到达间隔时间是指数分布的, 并且每单位时间顾客到达率为 λ, 则

$$\begin{aligned} p_0(t) &= \mathrm{P}\{\text{到达间隔时间} \geqslant t\} \\ &= 1 - \mathrm{P}\{\text{到达间隔时间} \leqslant t\} \\ &= 1 - (1 - \mathrm{e}^{-\lambda t}) \\ &= \mathrm{e}^{-\lambda t} \end{aligned}$$

对一充分小的时间区间 $h > 0$, 我们有

$$p_0(h) = \mathrm{e}^{-\lambda h} = 1 - \lambda h + \frac{(\lambda h)^2}{2!} - \cdots = 1 - \lambda h + 0(h^2)$$

指数分布基于假设: 在充分小的 $h > 0$ 期间, 最多有一个事件 (到达) 能够发生. 因此, 当 $h \to 0$,

$$p_1(h) = 1 - p_0(h) \approx \lambda h$$

这一结果表示, h 期间一次到达的概率直接与 h 成比例, 以到达率 λ 为比例常数.

在到达间隔时间服从平均值 $\frac{1}{\lambda}$ 的指数分布的条件下, 为了推导出某期间 t 内到达数目的分布, 定义

$$p_n(t) = t \text{ 期间有 } n \text{ 个到达的概率}$$

对充分小的 $h > 0$,

$$p_n(t+h) \approx \begin{cases} p_n(t)(1-\lambda h) + p_{n-1}(t)\lambda h, & n > 0 \\ p_0(t)(1-\lambda h), & n = 0 \end{cases}$$

第一个公式中, 如 t 时间内有 n 个到达而且在 h 时间内没有到达, 或者是在 t 期间有 $n-1$ 个到达, 在 h 期间有 1 个到达, 则在 $t+h$ 时间内有 n 个到达. 不可能有其他的组合, 因为根据指数分布, 在非常小的时间期间 h 内, 最多有一个到达. 因为每个个体到达与否是独立的, 因此可对等式的右端运用概率乘法法则. 对于第二个公式, 只有当 t 和 h 期间没有到达时, $t+h$ 期间才会有零个到达.

重新安排各项并取当 $h \to 0$ 的极限, 得到

$$p_n'(t) = \lim_{h \to 0} \frac{p_n(t+h) \quad p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad n > 0$$

$$p_0'(t) = \lim_{h \to 0} \frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t), \quad n = 0$$

其中 $p'_n(t)$ 是 $p_n(t)$ 关于 t 的一阶导数.

求解上述差分–微分方程, 得到

$$p_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, \quad n = 0, 1, 2, \cdots$$

这正是 t 期间平均有 $E\{n|t\} = \lambda t$ 个到达的**泊松分布** (Poisson distribution).

上面的结果说明, 若到达间隔时间服从平均值为 $\frac{1}{\lambda}$ 的指数分布, 则指定期间 t 内的到达数服从平均值 λt 的泊松分布. 反之亦然.

下表总结了给定单位时间到达率 λ 的指数分布与泊松分布之间的密切关系:

	指数分布	泊松分布
随机变量	相继到达之间的时间 t	指定 T 期间的到达数 n
取值范围	$t \geqslant 0$	$n = 0, 1, 2, \cdots$
密度函数	$f(t) = \lambda e^{-\lambda t}, t \geqslant 0$	$p_n(T) = \frac{(\lambda T)^n e^{-\lambda T}}{n!}, n = 0, 1, 2, \cdots$
平均值	$\frac{1}{\lambda}$ 时间单元	T 期间有 λT 个到达
累积概率	$P\{t \leqslant A\} = 1 - e^{-\lambda A}$	$p_{n \leqslant N}(T) = p_0(T) + p_1(T) + \cdots + p_N(T)$
$P\{A$ 期间无到达$\}$	$P\{t > A\} = e^{-\lambda A}$	$p_0(A) = e^{-\lambda A}$

例 12.4-1

某人口稀少州的出生率为每 12 分钟出生一个新生儿. 出生间隔时间服从指数分布. 求下列各值:

(a) 每年出生的平均数. (b) 任何一天内无新生儿出生的概率.

(c) 假设在 3 小时时间内的前 2 个小时已经发出了 40 份出生证明, 求这 3 小时内发出 50 份出生证明的概率.

计算每天的出生率为

$$\lambda = \frac{24 \times 60}{12} = 120\text{人/天}$$

该州每年出生人口数为

$$\lambda t = 120 \times 365 = 43\,800\text{人/年}$$

任意一天没有新生儿出生的概率可从泊松分布计算为

$$p_0(1) = \frac{(120 \times 1)^0 e^{-120 \times 1}}{0!} = e^{-120} = 0$$

计算这个概率的另外一种方法是, 因为任意一天内没有出生新生儿等价于说, 所以相继出生的间隔时间超过一天. 因此可以用指数分布来计算所需的概率

$$P\{t > 1\} = e^{-120} = 0$$

假设在 3 小时时间内的前 2 个小时已经发出了 40 份出生证明, 计算这 3 小时内发出 50 份出生证明的概率, 相当于 1 小时 $(=3-2)$ 内出生 $10(=50-40)$ 个新生儿, 因为出生数的分布是泊松的.

已知每小时出生数为 $\lambda=\frac{60}{12}=5$, 我们得出

$$p_{10}(1)=\frac{(5\times1)^{10}e^{-5\times1}}{10!}=0.018\ 13$$

Excel 程序

有关泊松分布以及所有排队公式的计算都比较麻烦, 要求有特殊的编程技巧来保证合适的计算精度. 可以用 Excel POISSON、POISSONDIST 和 EXPONDIST 函数来计算泊松分布和指数分布的单个概率和累积概率. 这些函数在 excelStatTables.xls 中也可以自动计算. 例如, 已知出生率为每小时 5 个婴儿, 要计算 0.5 小时内恰好有 10 个婴儿出生的概率, 我们在 F16 中输入 2.5, 在 J16 中输入 10, 就可以在 M16 中得出答案 0.000 216. O16 给出*最多*10 个婴儿出生的累积概率 (= 0.999 938). 为了求出生间隔时间小于等于 18 分钟的概率, 我们用指数分布程序, 在 F9 中输入 2.5, 在 J9 中输入 0.3, 则可在 O9 中得到答案 0.527 633.

TORA/Excel 程序

可以用 TORA(文件 toraEx12.4-1.txt) 或者 excelPoissonQ.xls 程序来自动计算所有有意义的 (TORA 中 $>10^{-5}$, Excel 中 $>10^{-7}$) 泊松概率. 在例 12.4-1 的纯生模型中, 数据输入如下:

Lambda	Mu	c	系统上限	输入源上限
5	0	无	无穷大	无穷大

注意对 Lambda 的输入为每小时出生 $\lambda t=5\times1=5$ 个.

习题 12.4A

***1.** 在例 12.4-1 中, 设工作人员通常等待直到累积了至少 5 个出生证明才把这些出生证明信息输入计算机. 求该工作人员每小时输入一批新信息的概率.

2. 一位艺术品收藏家平均一个月去一趟艺术品拍卖会, 每趟都能保证买回艺术品. 每两趟之间的间隔时间服从指数分布. 求以下各值:

(a) 3 个月没有购买的概率.　(b) 每年不超过 8 次购买的概率.

(c) 相继两趟的间隔时间超过 1 个月的概率.

3. 银行服务工作中, 到达率为每分钟 2 名顾客, 求下列各值:

(a) 5 分钟内平均到达数.　(b) 未来 0.5 分钟没有到达发生的概率.

(c) 未来 0.5 分钟至少一个到达发生的概率.

(d) 两个相继到达之间至少是 3 分钟的概率.

4. L&J 餐馆的到达间隔时间服从指数分布, 平均值为 5 分钟. 该餐馆从上午 11 点开始营业, 求下列值:

*(a) 已知到 11:05 A.M. 有 8 位顾客到达, 求到 11:12 A.M. 有 10 位顾客到达的概率.

(b) 已知上一位顾客 11:25 A.M. 到达, 求 11:28~11:33 A.M. 有新顾客到达的概率.

5. Springdale 公共图书馆按照泊松分布进一批新书, 平均每天进 25 本. 每层书架可以放置 100 本书. 求下列值:

(a) 每月 (30 天) 放置新书书架的平均层数.

(b) 已知一个书架有 5 层, 求每月需要 10 个以上书架的概率.

6. A 大学运行两条校内巴士线路: 红线和绿线. 红线服务北校区, 绿线服务南校区, 并由一个换乘站相连. 绿线巴士每 10 分钟 (按泊松分布) 随机到达换乘站, 红线巴士每 7 分钟也随机到达.

*(a) 5 分钟内两辆巴士同时停靠在换乘站的概率是多少?

(b) 宿舍靠近换乘站的某学生在 10 分钟内要上课, 这两辆巴士都能把该学生送到该教学楼. 乘车时间需要 5 分钟, 下车后他还要步行大约 3 分钟到达教室. 该学生能按时上课的概率是多少?

7. 证明 t 时间期间内的泊松分布的平均值和方差等于 λt, 其中 λ 为到达率.

8. 从纯生模型的差分–微分方程导出泊松分布. 提示: 一般微分方程

$$y' + a(t)y = b(t)$$

的解为

$$y = \mathrm{e}^{-\int a(t)\mathrm{d}t}\left\{\int b(t)\mathrm{e}^{\int a(t)\mathrm{d}t}\mathrm{d}t + c\right\} \quad (\text{其中 } c \text{ 为常数})$$

12.4.2 纯灭模型

纯灭模型中, 系统在 0 时刻开始时有 N 个顾客, 后面没有新顾客到达. 离开的发生率为每单位时间 μ 个顾客. 为了建立 t 时间单位后剩下 n 个顾客的概率 $p_n(t)$ 的差分–微分方程, 我们采用与纯生模型一样的方法 (见 12.4.1 节), 有

$$\begin{aligned}
p_N(t+h) &= p_N(t)(1-\mu h)\\
p_n(t+h) &= p_n(t)(1-\mu h) + p_{n+1}(t)\mu h, \quad 0 < n < N\\
p_0(t+h) &= p_0(t)(1) + p_1(t)\mu h
\end{aligned}$$

当 $h \to 0$, 得到

$$\begin{aligned}
p'_N(t) &= -\mu p_N(t)\\
p'_n(t) &= -\mu p_n(t) + \mu p_{n+1}(t), \quad 0 < n < N\\
p'_0(t) &= \mu p_1(t)
\end{aligned}$$

这组方程的解得到下面的**截尾泊松** (truncated Poisson) 分布:

$$p_n(t)=\frac{(\mu t)^{N-n}e^{-\mu t}}{(N-n)!},\quad n=1,2,\cdots,N$$

$$p_0(t)=1-\sum_{n=1}^{N}p_n(t)$$

例 12.4-2

某杂货店鲜花柜台每周初库存 18 打玫瑰花. 平均情况下, 鲜花柜台每天卖出 3 打 (一次一打), 但实际需求量服从泊松分布. 一旦库存水平剩下 5 打, 就再订货 18 打, 下周一送货. 由于鲜花商品的特性, 周末没有卖出的玫瑰花就要扔掉. 请求出下列各值:

(a) 该周内任何一天订货的概率.

(b) 周末扔掉玫瑰花的平均打数.

因为购买的发生率为每天 $\mu=3$ 打, t 日结束前订货的概率为

$$\begin{aligned}p_{n\leqslant 5}(t)&=p_0(t)+p_1(t)+\cdots+p_5(t)\\&=p_0(t)+\sum_{n=1}^{5}\frac{(3t)^{18-n}e^{-3t}}{(18-n)!},\quad t=1,2,\cdots,7\end{aligned}$$

我们最好用 excelPoissonQ.xls 或 TORA 来计算 $p_{n\leqslant 5}(t)$, TORA 的多情景功能对本例的计算更加方便. 对应于 $t=1,2,\cdots,7$ 纯灭模型相应的输入数据为

$$\text{Lambda}=0, \text{Mu}=3t, c=1, \text{系统上限}=18, \text{输入源上限}=18$$

注意到 t 必须要数值替换出来, 如文件 toraEx12.4-2.txt 所示.

输出结果如下:

t(星期几)	1	2	3	4	5	6	7
μt	3	6	9	12	15	18	21
$p_{n\leqslant 5}(t)$	0.000 0	0.008 8	0.124 2	0.424 0	0.732 4	0.908 3	0.975 5

周末 ($t=7$) 扔掉的玫瑰花平均打数为 $E\{n|t=7\}$. 为了计算这个值, 我们需要用到 $p_n(7), n=0,1,2,\cdots,18$, 这些值可以用提供的软件求出, 这样我们就得到

$$E\{n|t=7\}=\sum_{n=0}^{18}np_n(7)=0.664\approx 1\text{打}$$

习题 12.4B

1. 在例 12.4-2 中, 用 excelPoissonQ.xls 或 TORA 程序计算 $p_n(7), n=1,2,\cdots,18$, 然后手工验证从这些概率可求得 $E\{n|t=7\}=0.664$ 打.

2. 在例 12.4-2 中, 对于下列每个情况, 先写出代数解, 再用 excelPoissonQ.xls 或 TORA 程序给出数值答案.

*(a) 3 天后库存用光的概率. (b) 第二天末剩余的玫瑰花平均打数.

*(c) 已知第三天末最后一打被买走, 到第四天末有人买至少 1 打玫瑰花的概率.

(d) 已知一天前有人买花, 求下一次购买前的时间最多还有半天的概率.

(e) 第一天没有人买花的概率. (f) 周末不用订货的概率.

3. Springdale 高中管乐队正在一座新建的 400 个座位的礼堂进行爵士音乐会募捐演出. 当地企业按 10 张一组购票, 并将票款捐给青年会组织. 只在演出当天 4 小时内对企业售票, 订票过程服从平均值为每小时 10 次的泊松分布. 售票处关门以后剩下的票 (按组) 将在音乐会开始前 1 小时作为"临时票"按照折扣价卖出. 求

(a) 有可能买到临时票的概率. (b) 还有临时票的平均张数.

4. 每天早上, 一家小型机械加工厂的冰箱里都放上 2 箱 (每箱 24 罐) 软饮料, 供厂里的 10 名员工饮用. 这些员工在 8 小时工作日内 (8:00 A.M.~4:00 P.M.) 任何时间都可以喝饮料解渴, 且已知每个员工一天大约消费 4 罐, 但这一过程是完全随机的 (服从泊松分布). 那么一个员工中午 (午餐开始) 时就找不到饮料的概率是多少? 工厂关门前喝完所有饮料的概率是多少?

***5.** 某个一年级大学生每月家里给他 \$100 的存款, 用来支付杂费. 每次取款 \$20 会在这个月里随机发生, 取款的时间间隔服从指数分布, 平均值为 1 周. 求该学生在第 4 周末用完这个月杂费的概率.

6. 从仓库中按照泊松分布提取共 80 件库存产品, 提取率为每天 5 件, 求下列各项:

(a) 前 2 天提取 10 件库存的概率. (b) 第 4 天末库存用完的概率.

(c) 4 天期间提取库存产品的平均数.

7. 某机械加工厂刚好有 10 件备用零部件用来维修机器, 每 7 天补充一次库存, 保持 10 件零部件的库存水平. 机器出现故障的间隔时间服从指数分布, 平均值为 1 天. 求由于缺乏零部件导致机器 2 天不能工作的概率.

8. 对某商品的需求服从泊松分布, 平均每天 3 件. 最大库存水平为 25 件, 发生在每星期一, 刚好新的订货送到. 订货量的多少根据每周末星期六 (星期天不营业) 剩余的库存量决定. 求以下各项:

*(a) 每周平均订货量. *(b) 在星期五早上开门时发生缺货的概率.

(c) 每周订货量超过 10 件的概率.

9. 证明对应于纯灭模型中的截尾泊松分布, 离开间隔时间的分布服从平均值为 $\frac{1}{\mu}$ 时间单位的指数分布.

10. 从纯灭模型的差分–微分方程中, 用归纳法推导出截尾泊松分布. (注意: 见习题 12.4A 中第 8 题的提示.)

12.5 广义泊松排队模型

本节建立一个一般性的排队模型, 根据泊松分布的假设, 即到达间隔时间和服

务时间服从指数分布, 我们把到达和离开结合起来. 基于这一模型, 可以导出 12.6 节中的特殊泊松模型.

广义模型的建立是基于排队情形的长期行为, 或称为**平稳状态** (steady-state) 行为, 这种状态在系统经过了充分长时间的运行后得到的. 这种分析和系统初期运行期间所常见的**瞬时** (transient)(或称为热身) 行为完全不同. 本章不讨论瞬时行为的一个原因是由于对它的解析太复杂, 另一个原因是由于对大多数排队系统都是在平稳状态下来研究的.

广义模型假设, 到达率和离开率都是**与状态相关的** (state dependent), 也就是说, 它们都依赖于服务设施中的顾客数. 例如, 在高速公路收费口, 在高峰时间收费员通常就会提高收费速度. 另一个例子发生在商店, 假定收费机的数量一定, 随着坏了的机器数增加, 故障率就会减小 (因为只有正在工作的机器才可能发生新的故障).

定义

n = 系统中的顾客总数 (排队的加上正接受服务的)

λ_n = 已知系统中有 n 个顾客时的到达率

μ_n = 已知系统中有 n 个顾客时的离开率

p_n = 系统中有 n 个顾客的平稳状态概率

广义模型中 p_n 作为 λ_n 和 μ_n 的函数, 然后用这些概率来求出系统行为的度量指标, 如平均队长、平均等待时间以及设备平均利用率.

概率 p_n 可以用图 12.2 中的**转移率图** (transition-rate diagram) 来得到. 这个排队系统处在状态 n, 因为这时系统中的顾客数为 n. 根据 12.3 节中的解释, 在一个小时间区间 h 里多于一个事件发生的概率随着 $h \to 0$ 而趋于 0. 这就意味着, 对于 $n > 0$, 状态 n 只能变成两种可能的状态：当按照离开率 μ_n 离开时变成 $n-1$, 当按照到达率 λ_n 到达时变成 $n+1$. 状态 0 按照到达率 λ_0 到达时只能变成状态 1. 注意到假如系统为空时, 因为没有离开发生, μ_0 没有定义.

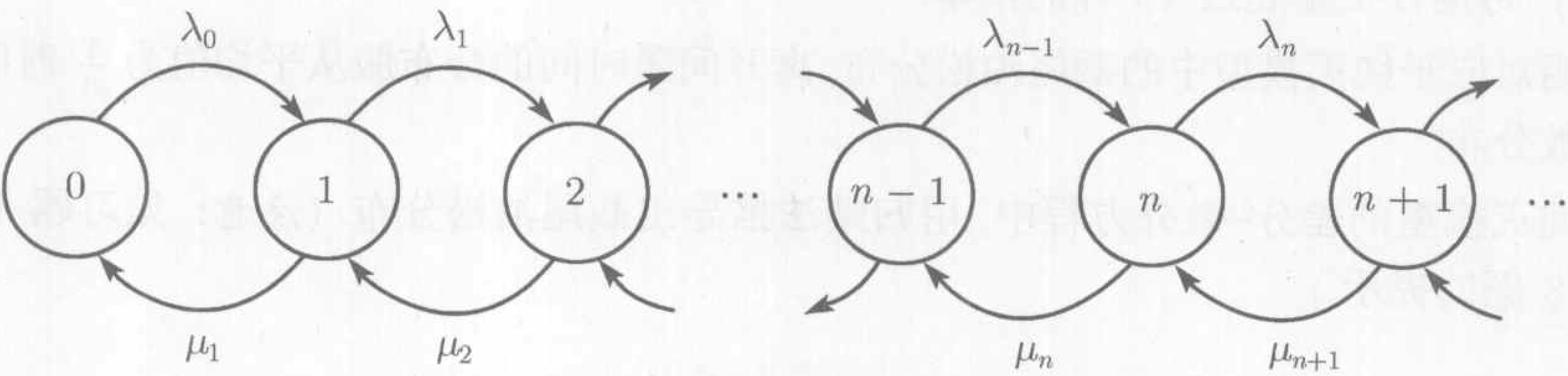


图 12.2　泊松队列转移图

在平稳状态条件下, 对于 $n > 0$, 流入和流出状态 n 的期望速率必相等. 根据

状态 n 只能变成状态 $(n-1)$ 和 $(n+1)$ 的事实, 我们得到

$$\text{流进状态 } n \text{ 的期望速率} = \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1}$$

类似地

$$\text{流出状态 } n \text{ 的期望率} = (\lambda_n + \mu_n)p_n$$

让这两个速率相等, 得到下面的**平衡方程** (balance equation):

$$\lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1} = (\lambda_n + \mu_n)p_n, \quad n = 1, 2, \cdots$$

从图 12.2 中, 对应于 $n = 0$ 的平衡方程为

$$\lambda_0 p_0 = \mu_1 p_1$$

从 p_0 开始递归求解平衡方程如下: 对于 $n = 0$, 有

$$p_1 = \left(\frac{\lambda_0}{\mu_1}\right) p_0$$

接下来, 对 $n = 1$, 有

$$\lambda_0 p_0 + \mu_2 p_2 = (\lambda_1 + \mu_1)p_1$$

用 $p_1 = (\frac{\lambda_0}{\mu_0})p_0$ 替换并化简, 得到 (请验证!)

$$p_2 = \left(\frac{\lambda_1\lambda_0}{\mu_2\mu_1}\right) p_0$$

一般地, 可用归纳法得到

$$p_n = \left(\frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1}\right) p_0, \quad n = 1, 2, \cdots$$

p_0 的值可以从等式 $\sum\limits_{n=0}^{\infty} p_n = 1$ 求出.

例 12.5-1

B&K 食品店有 3 个收款台. 经理根据店内顾客数量按照下列安排决定提供服务的收款台数:

店内顾客数	使用收款台数
$1 \sim 3$	1
$4 \sim 6$	2
6 人以上	3

顾客按照平均速度为每小时 10 位的泊松分布来到收款区. 每位顾客的平均收款时间为指数分布, 平均 12 分钟. 求 n 个顾客在收款区的平稳状态概率 p_n.

根据本问题的信息, 有

$$\lambda_n = \lambda = 10\text{个顾客/小时}, \quad n = 0, 1, \cdots$$

$$\mu_n = \begin{cases} \frac{60}{12} = 5\text{个顾客/小时}, & n = 0, 1, 2, 3 \\ 2 \times 5 = 10\text{个顾客/小时}, & n = 4, 5, 6 \\ 3 \times 5 = 15\text{个顾客/小时}, & n = 7, 8, \cdots \end{cases}$$

因此,

$$p_1 = \left(\tfrac{10}{5}\right) p_0 = 2p_0$$

$$p_2 = \left(\tfrac{10}{5}\right)^2 p_0 = 4p_0$$

$$p_3 = \left(\tfrac{10}{5}\right)^3 p_0 = 8p_0$$

$$p_4 = \left(\tfrac{10}{5}\right)^3 \left(\tfrac{10}{10}\right) p_0 = 8p_0$$

$$p_5 = \left(\tfrac{10}{5}\right)^3 \left(\tfrac{10}{10}\right)^2 p_0 = 8p_0$$

$$p_6 = \left(\tfrac{10}{5}\right)^3 \left(\tfrac{10}{10}\right)^3 p_0 = 8p_0$$

$$p_{n\geqslant 7} = \left(\tfrac{10}{5}\right)^3 \left(\tfrac{10}{10}\right)^3 \left(\tfrac{10}{15}\right)^{n-6} p_0 = 8\left(\tfrac{2}{3}\right)^{n-6} p_0$$

p_0 的值从下面的等式求出

$$p_0 + p_0\{2 + 4 + 8 + 8 + 8 + 8 + 8(\tfrac{2}{3}) + 8(\tfrac{2}{3})^2 + 8(\tfrac{2}{3})^3 + \cdots\} = 1$$

或等价地

$$p_0\{31 + 8[1 + \tfrac{2}{3} + (\tfrac{2}{3})^2 + \cdots]\} = 1$$

利用几何求级数

$$\sum_{i=0}^{\infty} x^i = \tfrac{1}{1-x}, \quad |x| < 1$$

得到

$$p_0\{31 + 8(\tfrac{1}{1-\frac{2}{3}})\} = 1$$

因此, $p_0 = \frac{1}{55}$.

知道了 p_0, 现在就能求出 p_n $(n > 0)$. 例如, 只有一个收款台开放的概率就可以计算为系统中最多有 3 个顾客的概率:

$$p_1 + p_2 + p_3 = (2 + 4 + 8)(\tfrac{1}{55}) \approx 0.255$$

可以用 p_n 来求出 B&K 食品店排队行为的度量指标, 例如

期望空闲收款台数 $= 3p_0+2(p_1+p_2+p_3)+1(p_4+p_5+p_6)+0(p_7+p_8+\cdots) = 1$个收款台

习题 12.5A

1. 在例 12.5-1 中, 求下列各项:

(a) 开放收款台个数的概率分布. (b) 繁忙收款台的平均数.

2. 在例 12.5-1 的 B&K 模型中, 假设收款区的到达间隔时间服从平均值为 5 分钟的指数分布, 并且每位顾客的收款时间也是指数分布的, 平均为 10 分钟. 进一步假设, B&K 将要增加第 4 个收款台, 这些收款台根据两个顾客的增幅而逐一增加开放. 求下列各值.

(a) 对所有 n, 求平稳状态概率 p_n. (b) 需要第 4 个收款台的概率.

(c) 平均空闲收款台数.

***3.** 在例 12.5-1 的 B&K 模型中, 假设所有 3 个收款台总是开放, 运作方式是顾客将去第一个空闲的收款台. 求下列各值:

(a) 所有 3 个收款台都在使用的概率. (b)顾客到达不用等待的概率.

4. Springdale 第一银行提供一台单车道免下车 ATM 取款机. 汽车按照泊松分布每小时 12 辆的到达率到达. 每辆车完成 ATM 业务所需要的时间服从平均 6 分钟的指数分布. 车道总共可允许 10 辆客车驶入. 一旦车道占满了, 其他来到的车辆就要去另一个分行寻求服务了. 求下列各项:

(a) 一辆来到的车由于车道占满而不能使用 ATM 机的概率.

(b) 一辆车到来后不能立即使用 ATM 机的概率.

(c) 车道里的平均车辆数.

5. 你听过有人老是说“那个地方太拥挤了, 没有人再会去那儿了”这样矛盾的话吗? 这句话可以解释为, 随着寻求服务的顾客人数的增加, “暂时不加入”的机会也在增加. 对这种情形建模的一种可能的依据是说, 随着系统中顾客人数的增长, 系统的到达率就会下降. 举例来说, 我们考虑 M&M 台球俱乐部的简化情形, 顾客们成组 (每组 2 人) 来打台球. 正常的到达率是每小时 6 组 (12 人). 但是一旦台球厅里打球的人超过 8 组时, 到达率就降至每小时来 5 组. 假定这一到达过程服从泊松分布. 每组人玩平均 30 分钟的指数分布时间. 台球厅总共有 5 个球台, 任何时候可供不超过 12 组人玩. 求下列各项:

(a) 顾客们暂不加入排队的概率. (b) 所有球台都在使用的概率.

(c) 在使用中的球台的平均数. (d) 等待球台的人的平均组数.

***6.** 某理发店一次服务一位顾客, 为等待的顾客准备了 3 把椅子. 假如这些位置都满了, 顾客就到别的地方去了. 顾客的到达服从泊松分布, 平均每小时 4 位. 每位顾客理发的时间服从平均值为 15 分钟的指数分布. 求下列各值:

(a) 平稳状态概率. (b) 店中顾客数的期望值.

(c) 由于满员使得顾客到别的地方去的概率.

7. 考虑一个服务台的排队情形, 其中到达率和服务率分别为

$$\lambda_n = 10 - n, \quad n = 0, 1, 2, 3$$

$$\mu_n = \frac{n}{2} + 5, \quad n = 1, 2, 3, 4$$

这个情形等价于随着系统中 n 的增加来减小到达率和增加服务率.

(a) 建立一个状态转移图并确定该系统的平衡方程.

(b) 求平稳状态的概率.

8. 考虑单一排队模型, 其中只允许一个顾客进入系统. 到达并发现服务设施忙的顾客不会再返回. 假设到达分布是泊松的, 平均每单位时间到达人数为 λ, 服务时间服从平均值为 $\frac{1}{\mu}$ 的指数分布.

(a) 建立状态转移图并求平衡方程.　(b) 求平稳状态的概率.

(c) 求系统中平均顾客数.

9. 用归纳法证明得出下面广义模型的一般解. 考虑

$$p_k = \prod_{i=0}^{k-1}\left(\frac{\lambda_i}{\mu_{i+1}}\right)p_0,\quad k=0,1,2,\cdots$$

我们在包含 p_n, p_{n-1}, p_{n-2} 的一般差分方程中代换 p_{n-1} 和 p_{n-2}, 以得到所要的 p_n 的表达式. 证明这一方法的正确性.

12.6　特殊泊松队列

图 12.3 显示带有 c 个并行服务台的特殊泊松排队情形. 由第一个空闲的服务台从队列中选择一个等待顾客开始服务. 系统的到达率为每单位时间 λ 个顾客. 所有服务台都一样, 意味着每个服务台的服务率都是每单位时间 μ 个顾客. *系统中的顾客数定义为正接受服务的顾客加上在队列中等待服务的顾客*.

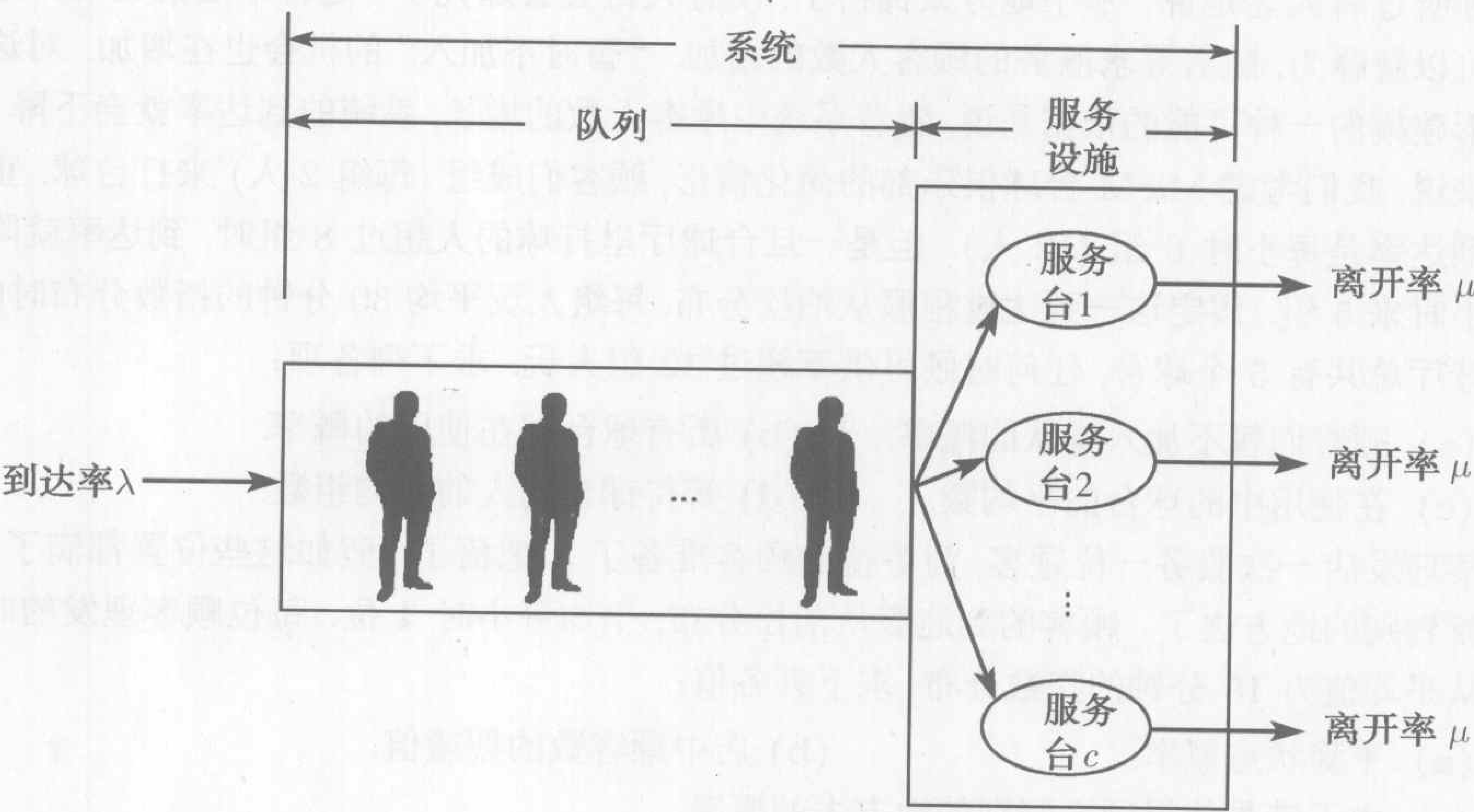


图 12.3　有 c 个并行服务台的排队系统的图形表示

为了用方便的符号表示图 12.3 这种排队情形的特性, 采用下面的格式

$$(a/b/c):(d/e/f)$$

其中,

$a=$ 到达分布

$b=$ 离开 (服务时间) 的分布

$c=$ 并行服务台数目($=1,2,\cdots,\infty$)

$d=$ 排队规则

$e=$ 系统中 (队列中的加上正在接受服务的) 最大允许数 (有限或无限多个)

$f=$ 顾客输入源的多少 (有限或无限多个)

表示到达和离开分布 (符号 a 和 b) 的标准记号有

$M=$ 马尔可夫 (或泊松) 到达或离开分布 (等价于指数到达间隔时间或服务时间分布)

$D=$ 常数 (确定性) 时间

$E_k=$ 埃尔朗或 Γ 时间分布 (等价于独立指数分布之和)

$GI=$ 到达间隔时间的一般性 (通用) 分布

$G=$ 服务时间的一般性 (通用) 分布

排队规则的记号 (符号 d) 包括

$FCFS=$ 先到先服务

$LCFS=$ 后到先服务

$SIRO=$ 随机次序服务

$GD=$ 一般规则 (即任何一种规则)

举例说明如何使用这些记号. 模型 $(M/D/10):(GD/20/\infty)$ 采用泊松到达 (或指数到达间隔时间), 常数服务时间, 10 个并行服务台. 排队规则是一般规则, 整个系统中只限于有 20 个顾客, 顾客输入源的大小是无限的.

根据历史记载, 这种记号的前 3 个要素 $(a/b/c)$ 是 1953 年由 D.G. 肯德尔提出的, 所以通常称为**肯德尔记号** (Kendall notation). 1966 年, A.M. 李在这个记号中又加了符号 d 和 e, 该作者 1968 年又加上了最后一个符号 f.

在详细介绍特殊泊松队列之前, 先说明如何从 12.5 节给出的平稳状态概率 p_n 得到广义排队系统的平稳状态度量指标.

12.6.1 队列行为的平稳状态度量

最常用的队列行为度量指标有

$L_s=$ 系统中顾客的期望数

$L_q=$ 队列中顾客的期望数

$W_s=$ 系统中的期望等待时间

$W_q=$ 队列中的期望等待时间

$\bar{c}=$ 繁忙服务台的期望数

前面说过, 系统包括队列加上服务设施.

现在说明这些指标如何 (直接或间接地) 从系统中状态 n 的平稳状态概率 p_n 来得到:

$$L_s = \sum_{n=1}^{\infty} np_n$$

$$L_q = \sum_{n=c+1}^{\infty} (n-c)p_n$$

L_s 与 W_s(以及 L_q 与 W_q) 之间的关系称为**Little 公式** (Little's formula), 具体关系为

$$L_s = \lambda_{\text{eff}} W_s$$

$$L_q = \lambda_{\text{eff}} W_q$$

这些关系在相当一般的条件下成立. 参数 λ_{eff} 是系统的有效到达率. 当所有到达的顾客都可能加入系统时, 它就等于到达率 λ. 否则的话, 如果由于系统满了 (例如停车场), 有些顾客不能加入, 则 $\lambda_{\text{eff}} < \lambda$. 后面将说明如何求出 λ_{eff}.

W_s 和 W_q 之间也存在直接关系. 由定义

$$\text{期望系统等待时间} = \text{期望队列等待时间} + \text{期望服务时间}$$

可以写成

$$W_s = W_q + \frac{1}{\mu}$$

可以通过将这一公式两边乘以 λ_{eff}, 建立 L_s 与 L_q 的关系. 结合 Little 公式一起, 得到

$$L_s = L_q + \frac{\lambda_{\text{eff}}}{\mu}$$

根据定义, 系统中的平均数 L_s 与队列中的平均数 L_q 之差等于繁忙服务台的平均数 $\bar{c}$. 因此有

$$\bar{c} = L_s - L_q = \frac{\lambda_{\text{eff}}}{\mu}$$

就得到了

$$\text{设施利用率} = \frac{\bar{c}}{c}$$

例 12.6-1

Ozark 学院来访者的停车位只限于 5 个, 使用这些停车位的车辆按照泊松分布到达, 每小时平均到达 6 辆车. 停车时间服从均值为 30 分钟的指数分布. 到达后找不到空车位的来访者可以在停车场临时车位等待, 直到有停着的车离开. 临时车位

只能放 3 辆车. 其他停不了也找不到临时等待车位的车辆必须去别的地方. 求下列各项:

(a) 系统中有 n 辆车的概率 p_n.

(b) 实际使用停车场的车辆的有效到达率.

(c) 车场平均车辆数.

(d) 一辆车在停车场内等待停车位的平均时间.

(e) 占据停车位的平均车辆数.

(f) 停车场的平均利用率.

首先注意到, 一个停车位就是一个服务台, 这样, 系统共有 $c=5$ 个并行服务台. 另外, 系统的最大容量为 $5+3=8$ 辆车.

概率 p_n 可以按 12.5 节广义模型的特殊情况来求出, 用公式

$$\lambda_n = 6 \text{ 辆车/小时}, \quad n=0,1,2,\cdots,8$$

$$\mu_n = \begin{cases} n\left(\frac{60}{30}\right) = 2n \text{ 辆车/小时}, & n=1,2,3,4,5 \\ 5\left(\frac{60}{30}\right) = 10 \text{ 辆车/小时}, & n=6,7,8 \end{cases}$$

根据 12.5 节, 得到

$$p_n = \begin{cases} \dfrac{3^n}{n!}p_0, & n=1,2,3,4,5 \\ \dfrac{3^n}{5!5^{n-5}}p_0, & n=6,7,8 \end{cases}$$

p_0 的计算可通过把 $p_n, n=1,2,\cdots,8$ 代入到下面的方程中

$$p_0+p_1+\cdots+p_8=1$$

即

$$p_0+p_0\left(\frac{3}{1!}+\frac{3^2}{2!}+\frac{3^3}{3!}+\frac{3^4}{4!}+\frac{3^5}{5!}+\frac{3^6}{5!5}+\frac{3^7}{5!5^2}+\frac{3^8}{5!5^3}\right)=1$$

由此得到 $p_0=0.048\ 12$(请验证!). 有了 p_0, 现在可以计算 p_1 到 p_8 如下:

n	1	2	3	4	5	6	7	8
p_n	0.144 36	0.216 54	0.216 54	0.162 40	0.097 44	0.058 47	0.035 08	0.021 05

有效到达率 λ_{eff} 可以通过观察图 12.4 的示意图算出, 其中顾客从输入源以每小时 λ 辆车的到达率到达. 一辆到达的车可以按照到达率 λ_{eff} 进入停车场或者以 λ_{lost} 离开, 这意味着 $\lambda=\lambda_{\text{eff}}+\lambda_{\text{lost}}$. 假如有 8 辆车已经在停车场内, 则来到的车就不能再进入停车场. 这就意味着不能进入停车场车辆的概率是 p_8. 因此

$$\lambda_{\text{lost}}=\lambda p_8=6\times 0.021\ 05=0.126\ 3 \text{ 辆车/小时}$$

$$\lambda_{\text{eff}}=\lambda-\lambda_{\text{lost}}=6-0.126\ 3=5.873\ 7 \text{ 辆车/小时}$$

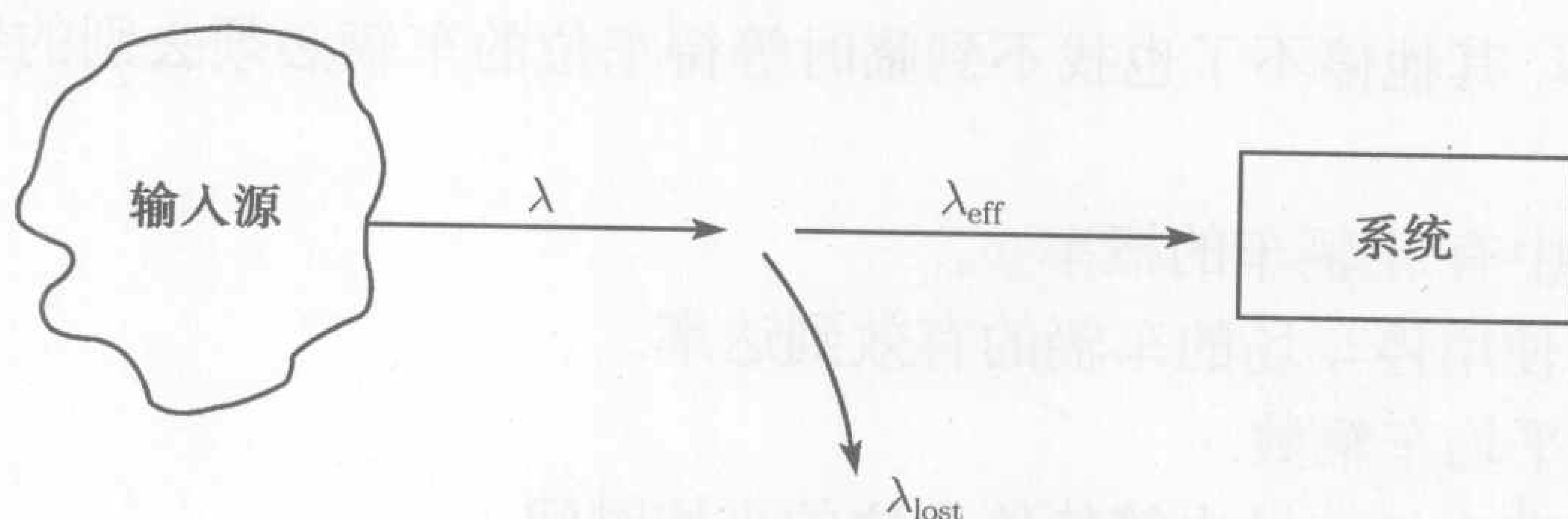


图 12.4　λ、λ_{eff} 及 λ_{lost} 之间的关系

停车场内 (等待或占用了停车位) 车辆的平均数, 等于系统内车辆的平均数 L_s. 我们可以用 p_n 来计算出 L_s:

$$L_s = 0p_0 + 1p_1 + \cdots + 8p_8 = 3.128\ 6 \text{ 辆车}$$

在临时车位等待的车辆实际上就是队列中的车辆. 因此, 等待找到车位的等待时间就是 W_q. 为确定 W_q 用

$$W_q = W_s - \frac{1}{\mu}$$

因此

$$W_s = \frac{L_s}{\lambda_{\text{eff}}} = \frac{3.128\ 6}{5.873\ 7} = 0.532\ 65 \text{ 小时}$$

$$W_q = 0.532\ 65 - \frac{1}{2} = 0.032\ 65 \text{ 小时}$$

占用了车位的平均车辆数就与繁忙服务台的平均数相等,

$$\bar{c} = L_s - L_q = \frac{\lambda_{\text{eff}}}{\mu} = \frac{5.873\ 7}{2} = 2.936\ 8 \text{ 个车位}$$

从 $\bar{c}$ 得到

$$\text{停车场利用率} = \frac{\bar{c}}{c} = \frac{2.936\ 8}{5} = 0.587\ 36$$

习题 12.6A

1. 在例 12.6-1 中进行下列计算:
   *(a) 直接用公式 $\sum_{n=c+1}^{\infty}(n-c)p_n$ 计算 L_q.　　(b) 用 L_q 计算 W_s.
   *(c) 计算 8 小时内不能进入停车场的车辆平均数.
   *(d) 证明平均空车位数 $c-(L_s-L_q)$ 等于 $\sum_{n=0}^{c-1}(c-n)p_n$.
2. 利用下列数据求解例 12.6-1: 停车位数 =6, 临时车位数 =4, $\lambda=$ 每小时 10 辆车, 平均泊车时间 = 45 分钟.

12.6.2 单服务台模型

本节介绍单一服务台情况 ($c = 1$) 的两个模型. 第 1 个模型对系统中的最大顾客数没有限制, 第 2 个模型假定有限的系统上限. 这两个模型都假定有无限多个顾客来源. 到达率为每单位时间 λ 个顾客, 服务率为每单位时间服务 μ 个顾客.

这两个模型的结果 (也是 12.6 节中后续所有模型的结果) 是从 12.5 节广义模型结果的特殊情况下得到的.

我们用肯德尔记号来总结每种情形下的特征. 因为 12.5 节中的 p_n 以及 12.6.1 节中所有性能度量指标都与具体的排队规则没有任何关系, 所以在记法上我们用了 GD(一般排队规则).

$(M/M/1)$:$(GD/\infty/\infty)$ 利用广义模型的记法, 有

$$\left.\begin{aligned}\lambda_n = \lambda \\ \mu_n = \mu\end{aligned}\right\}, \quad n = 0, 1, 2, \cdots$$

并且, $\lambda_{\text{eff}} = \lambda, \lambda_{\text{lost}} = 0$, 因为所有的到达顾客都能加入到系统.

令 $\rho = \frac{\lambda}{\mu}$, 则广义模型中 p_n 的表达式就简化成

$$p_n = \rho^n p_0, \quad n = 0, 1, 2, \cdots$$

为了求 p_0 的值, 用等式

$$p_0(1 + \rho + \rho^2 + \cdots) = 1$$

设 $\rho < 1$, 几何级数将有有限和 $\frac{1}{1-\rho}$, 因此

$$p_0 = 1 - \rho, \quad 其中 \quad \rho < 1$$

所以 p_n 的一般公式由下面的几何分布给出:

$$p_n = (1 - \rho)\rho^n, \quad n = 1, 2, \cdots (\rho < 1)$$

p_n 的数学推导将用到条件 $\rho < 1$ 或 $\lambda < \mu$. 若 $\lambda \geqslant \mu$, 则几何级数发散, 平稳状态概率 p_n 不存在. 这个结果有着直观意义, 因为除非服务率大于到达率, 否则队列长度将会不断增长, 不可能达到平稳状态.

排队系统的性能度量指标 L_s 可以按下面方式得到:

$$\begin{aligned}L_s = \sum_{n=0}^{\infty} n p_n &= \sum_{n=0}^{\infty} n(1-\rho)\rho^n \\ &= (1-\rho)\rho \frac{\mathrm{d}}{\mathrm{d}\rho} \sum_{n=0}^{\infty} \rho^n \\ &= (1-\rho)\rho \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{1}{1-\rho}\right) = \frac{\rho}{1-\rho}\end{aligned}$$

因为对于本情形 $\lambda_{\text{eff}} = \lambda$, 剩下的系统性能度量指标用 12.6.1 节中的关系来计算. 因此有

$$W_s = \frac{L_s}{\lambda} = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu-\lambda}$$

$$W_q = W_s - \frac{1}{\mu} = \frac{\rho}{\mu(1-\rho)}$$

$$L_q = \lambda W_q = \frac{\rho^2}{1-\rho}$$

$$\bar{c} = L_s - L_q = \rho$$

例 12.6-2

Automata 洗车房只运行一个清洗位. 车辆按照泊松分布到达, 平均每小时 4 辆车, 如果清洗位忙, 则到达的车辆等在洗车房的停车场. 一辆车的清洗时间服从指数分布, 平均值为 10 分钟. 不能停进停车场的车辆可在洗车房旁的路边等待, 这意味着从实际上来讲, 系统的规模是没有限制的. 洗车房经理想要确定出停车位要多少合适.

对于这一情形, 有 $\lambda = 4$ 辆车/小时, $\mu = \frac{60}{10} = 6$ 辆车/小时. 因为 $\rho = \frac{\lambda}{\mu} < 1$. 系统可以按平稳状态运行.

这个模型的 TORA 或 excelPoissonQ.xls 程序输入是

Lambda	Mu	c	系统上限	顾客源上限
4	6	1	无限	无限

模型输出如图 12.5 所示. 队列中等待车辆平均数 L_q 为 1.33 辆车.

一般地, 我们不建议只用 L_s 来计算停车位数, 因为从某种意义上, 停车位设计代表了要保证最大可能的队列长度. 例如, 更加可信的是, 停车场的设计要使来到的车在至少 90% 的时候都能找到停车位. 为了做到这一点, 我们令 S 表示停车位数, 有 S 个停车位就等价于系统中有 S+1 个位置 (队列加上洗车位). 假如系统中最多有 S 辆车, 则到达的车辆 90% 的时候都能找到位置. 这个条件等价于下面的概率叙述:

$$p_0 + p_1 + \cdots + p_S \geqslant 0.9$$

从图 12.5 中看出, 对于 $n = 5$, 累计的 p_n 为 0.912 21. 这意味着对于 $S \geqslant 5$ 个停车位, 上述条件满足.

车位数 S 也可以用 p_n 的数学定义来求出, 即

$$(1-\rho)(1+\rho+\rho^2+\cdots+\rho^S) \geqslant 0.9$$

上述截尾几何级数的和等于 $\frac{1-\rho^{S+1}}{1-\rho}$, 因此条件简化为

$$(1-\rho^{S+1}) \geqslant 0.9$$

简化上述不等式得到

$$\rho^{S+1} \leqslant 0.1$$

不等式两边取对数 [并注意到 $\ln(x) < 0, 0 < x < 1$, 不等式变号], 就得到

$$S \geqslant \frac{\ln(0.1)}{\ln(\frac{4}{6})} - 1 = 4.679 \approx 5$$

```
Scenario1:   (M/M/1):(GD/infinity/infinity)

Lambda =        4.00000                      Mu =         6.00000
Lambda eff = 4.00000                         Rho/c =      0.66667
Ls =          2.00000                        Lq =          1.33333
Ws =       0.50000                           Wq =         0.33333
```

n	Probability pn	Cumulative Pn	n	Probability pn	Cumulative Pn
0	0.33333	0.33333	13	0.00171	0.99657
1	0.22222	0.55556	14	0.00114	0.99772
2	0.14815	0.70370	15	0.00076	0.99848
3	0.09877	0.80247	16	0.00051	0.99899
4	0.06584	0.86831	17	0.00034	0.99932
5	0.04390	0.91221	18	0.00023	0.99955
6	0.02926	0.94147	19	0.00015	0.99970
7	0.01951	0.96098	20	0.00010	0.99980
8	0.01301	0.97399	21	0.00007	0.99987
9	0.00867	0.98266	22	0.00004	0.99991
10	0.00578	0.98844	23	0.00003	0.99994
11	0.00385	0.99229	24	0.00002	0.99996
12	0.00257	0.99486	25	0.00001	0.99997

图 12.5 例 12.6-2 的 TORA 输出 (文件 toraEx12.6-2.txt)

习题 12.6B

1. 在例 12.6-2 中, 求下列各项:
   (a) 清洗位的利用百分比.
   (b) 到达车辆在进入清洗位之前必须在停车场等待的概率.
   (c) 假如有 7 个停车位, 求到达车辆必能找到空停车位的概率.
   (d) 应提供多少停车位, 以便让到达车辆在 99% 的时候都能找到停车位?

*2. John Macko 是 Ozark 大学的一名学生. 他靠打零工来补充收入. 工作要求平均每 5 天来一次, 但是不同工作要求之间的时间服从指数分布. 完成一项工作的时间也服从指数分布, 平均为 4 天.
   (a) John 失去工作的概率是多少?
   (b) 若每一项工作中 John 能挣大约 \$50, 他的平均月收入是多少?
   (c) 如果在学期末, John 决定按每项工作 \$40 把未完成的工作转给别人, 平均他要支付多少钱?

3. 多年以来, Fayetteville 警署的 Columbo 警探在侦破每件刑事案件中成绩斐然. 破获刑事案件只是个时间问题. Columbo 承认, 侦破每个案件的时间是“完全随机的”, 但平均来讲, 每次调查大约需要 1.5 个星期. 在安定的 Fayetteville 镇, 刑事案件并不多. 它们随机地发生, 每个月 (4 周) 平均一起. Columbo 警探要求助手分担案件侦破工作. 请分析 Columbo 的话, 特别从下面的角度来进行分析:

(a) 等待调查的平均案件数. (b) 警探忙于案件的时间百分率.

(c) 需要侦破一宗案件的平均时间.

4. 来到林肯隧道收费站的车辆服从泊松分布, 平均每小时 90 辆车. 通过收费口的时间服从指数分布, 平均为 38 秒. 司机们抱怨等待时间过长, 假如下面两个条件满足的话, 管理部门愿意通过安装自动收费设备来把平均通过时间降低到 30 秒: (1) 现有系统中等待车辆的平均数超过 5 辆; (2) 安装了新的收费设备后, 收费口空闲时间的百分率不超过 10%. 安装新设备是否值得?

***5.** 某快餐店有一个免下车服务窗口, 到达车辆服从泊松分布, 平均到达率为每 5 分钟 2 辆车. 服务窗口前面的位置可供最多 10 辆车驶入, 包括正在接受服务的 1 辆车. 其他车辆必要时可等待在停车位以外. 对每个顾客的服务时间是指数分布的, 平均 1.5 分钟. 求下列各项:

(a) 设施空闲的概率. (b) 等待服务的期望顾客数.

(c) 一顾客到达窗口订餐之前的期望等待时间.

(d) 等待队列超出 10 个车位的概率.

6. 顾客来到某个单窗口免下车服务银行, 到达时间服从泊松分布, 平均每小时 10 个顾客. 每个顾客的服务时间服从指数分布, 平均 5 分钟. 该窗口前有 3 个车位, 包括正在接受服务的车辆, 其他到达车辆在这 3 个车位以外等待.

(a) 一辆到达的车辆能驶入这 3 个车位其中之一的概率是多少?

(b) 一辆到达的车辆要等在这 3 个指定车位以外的概率是多少?

(c) 一顾客到达窗口开始服务之前的期望等待时间有多长?

*(d) 窗口前应提供多少车位 (包括正接受服务的车辆位置), 以便让到达的车辆在 90% 的时候都能找到车位?

7. 在 $(M/M/1):(GD/\infty/\infty)$ 中, 为什么 L_s 一般不等于 L_q+1, 请给出合理的解释. 等式通常在什么条件下成立.

8. 对于 $(M/M/1):(GD/\infty/\infty)$, 利用基本定义 $\sum_{n=2}^{\infty}(n-1)p_n$ 导出 L_q 的表达式.

9. 对于 $(M/M/1):(GD/\infty/\infty)$, 说明

(a) 已知队列非空, 队列中期望顾客数 $=\frac{1}{1-\rho}$.

(b) 对那些必须等待的顾客, 在队列中的期望等待时间 $=\frac{1}{1-\rho}$.

$(M/M/1)$:$(GD/\infty/\infty)$ 的等待时间分布(Waiting Time Distribution)① 12.5 节的广义模型中所得到的 p_n 与排队规则完全无关. 这意味着, 排队系统性能的平均度量指标 (W_s, W_q, L_s, L_q) 可用于所有的排队规则.

虽然平均等待时间与排队规则无关, 但它的概率密度函数却不然. 我们通过从基于 $FCFS$ 规则的 $(M/M/1)$ 模型导出等待时间的分布来说明这一点.

---

① 可选择阅读.

令 τ 为刚刚到达的一位顾客必须在系统中的时间量 (即, 直到服务完成的时间). 根据 $FCFS$ 规则, 假如一个刚到达的顾客前面还有 n 个顾客在系统中, 则

$$\tau = t_1' + t_2 + \cdots + t_{n+1}$$

其中 t_1' 是当前在接受服务的顾客所需要的完成服务的时间, $t_2, t_3, \cdots, t_n$ 为队列中的 $n-1$ 个顾客的服务时间. 时间 t_{n+1} 表示刚刚到来的顾客的服务时间.

定义 $w(\tau|n+1)$ 为已知系统中有 n 个顾客在刚刚到达的顾客之前的条件下, τ 的条件密度函数. 因为服务时间的分布是指数的, 指数分布的遗忘性质 (12.3 节) 告诉我们, t_1' 也是指数的, 服从相同的分布. 因此, τ 等于 $(n+1)$ 个同分布的独立指数随机变量之和. 由概率论知, $w(\tau|n+1)$ 服从带有参数 μ 和 $(n+1)$ 的 Γ 分布. 因此有

$$\begin{aligned}
w(\tau) &= \sum_{n=0}^{\infty} w(\tau|n+1)p_n \\
&= \sum_{n=0}^{\infty} \frac{\mu(\mu\tau)^n \mathrm{e}^{-\mu\tau}}{n!}(1-\rho)\rho^n \\
&= (1-\rho)\mu\mathrm{e}^{-\mu\tau} \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!} \\
&= (1-\rho)\mu\mathrm{e}^{-\mu\tau}\mathrm{e}^{\lambda\tau} \\
&= (\mu-\lambda)\mathrm{e}^{-(\mu-\lambda)\tau}, \quad \tau > 0
\end{aligned}$$

因此, $w(\tau)$ 为一指数分布, 其平均值为 $W_s = \frac{1}{\mu-\lambda}$.

例 12.6-3

在例 12.6-2 的洗车房模型中, 我们可以合理地假定, 这项服务是根据 $FCFS$ 规则进行的. 评价用 W_s 作为估计系统中等待时间的可靠性.

回答这个问题的一种方法是, 估计顾客中等待时间超过 W_s 的比例. 注意到 $W_s = \frac{1}{(\mu-\lambda)}$, 我们得到

$$\begin{aligned}
\mathrm{P}\{\tau > W_s\} &= 1 - \int_0^{W_s} w(\tau)\mathrm{d}\tau \\
&= \mathrm{e}^{-(\mu-\lambda)W_s} = \mathrm{e}^{-1} = 0.368
\end{aligned}$$

因此在 $FCFS$ 规则下, 大约有 37% 的顾客的等待时间比 W_s 要长. 这似乎有些过长了, 特别是因为现在洗车房的 W_s 已经很高了 (= 0.5 小时). 我们注意到, 所计算的概率 ($= \mathrm{e}^{-1} \approx 0.368$) 与任何 $(M/M/1):(FCFS/\infty/\infty)$ 的到达率 λ 和服务率 μ 都无关, 这意味着它的值不能减少. 这样, 如果我们根据平均值 W_s 来设计系统的话, 则我们应该期望有 36.8% 的顾客的等待时间长于平均等待时间.

这种情形可以用两种方法加以改善: (1) 我们可以增加服务率 μ, 把 W_s 的值降到一个可以接受的水平; (2) 我们可以选择服务率使得等待时间超过某个指定值 (如, 10 分钟) 的概率保持在一个相对小的百分率以下 (例如, 10%). 第 1 种方法等价于找到一个 μ, 使得 $W_s < \overline{T}$; 第 2 种方法是要通过解不等式 $\mathrm{P}\{\tau > \overline{T}\} < \alpha$, 求得 μ, 其中 $\overline{T}$ 和 α 由分析人员指定.

习题 12.6C

***1.** 在习题 12.6B 的第 3 题中, 求 Columbo 警探要花一周以上时间破一宗刑事案件的概率.

2. 在例 12.6-3 中, 计算下列各项:

(a) 系统中等待时间 τ 的标准差.

(b) 系统中等待时间围绕平均值变化幅度为标准差一半的概率.

3. 在例 12.6-3 中, 求满足条件 $W_s < 10$ 分钟的服务率 μ.

4. 在例 12.6-3 中, 求满足条件 $\mathrm{P}\{\tau > 10\text{分钟}\} < 0.1$ 的服务率 μ.

***5.** 考虑习题 12.6B 中的第 5 题. 为了吸引更多的业务, 该餐厅老板为等待时间超过 5 分钟的顾客提供免费软饮料. 已知一份饮料价值 50 美分, 假定餐厅每天营业 12 个小时, 每天提供免费软饮料要花掉多少钱?

6. 证明对于 $(M/M/1):(FCFS/\infty/\infty)$ 模型, 队列中等待时间的分布为

$$w_q(t) = \begin{cases} 1-\rho, & t=0 \\ \mu\rho(1-\rho)\mathrm{e}^{-(\mu-\lambda)t}, & t>0 \end{cases}$$

然后从 $w_q(t)$ 中求出 W_q.

$(M/M/1)$:$(GD/N/\infty)$ 这个模型和 $(M/M/1):(GD/\infty/\infty)$ 不同, 它对系统中的顾客数 (最大队列长度 $=N-1$) 有一个上限 N. 例如, 一台机器的生产过程有有限个缓冲区, 以及快餐店的单车道免下车窗口.

当系统中的顾客数达到 N 时, 不允许再有到达, 因此有

$$\lambda_n = \begin{cases} \lambda, & n=0,1,\cdots,N-1 \\ 0, & n=N, \quad N+1 \end{cases}$$

$$\mu_n = \mu, \quad n=0,1,\cdots$$

利用 $\rho = \frac{\lambda}{\mu}$, 12.5 节的广义模型给出

$$p_n = \begin{cases} \rho^n p_0, & n \leqslant N \\ 0, & n > N \end{cases}$$

可以从等式 $\sum\limits_{n=0}^{\infty} p_n = 1$ 求出 p_0 的值, 这将得到

$$p_0(1+\rho+\rho^2+\cdots+\rho^N) = 1$$

或

$$p_0 = \begin{cases} \dfrac{(1-\rho)}{1-\rho^{N+1}}, & \rho \neq 1 \\ \dfrac{1}{N+1}, & \rho = 1 \end{cases}$$

因此有

$$p_n = \begin{cases} \dfrac{(1-\rho)\rho^n}{1-\rho^{N+1}}, & \rho \neq 1 \\ \dfrac{1}{N+1}, & \rho = 1 \end{cases}, \quad n = 0, 1, \cdots, N$$

在这个模型中, $\rho = \frac{\lambda}{\mu}$ 的值不需要小于 1, 因为系统的到达受到系统顾客上限 N 的控制. 这就意味着在这种情况下, 重要的是到达率 λ_{eff} 而不是 λ. 由于当系统中有了 N 个顾客, 再来的顾客就不再进入系统, 如图 12.4 所示,

$$\lambda_{\text{lost}} = \lambda p_N$$

$$\lambda_{\text{eff}} = \lambda - \lambda_{\text{lost}} = \lambda(1 - p_N)$$

在这种情况下, $\lambda_{\text{eff}} < \mu$.

系统中的期望顾客数可计算为

$$\begin{aligned} L_s &= \sum_{n=0}^{N} n p_n \\ &= \frac{1-\rho}{1-\rho^{N+1}} \sum_{n=0}^{N} n\rho^n \\ &= \left(\frac{1-\rho}{1-\rho^{N+1}}\right) \rho \frac{d}{d\rho} \sum_{n=0}^{N} \rho^n \\ &= \frac{(1-\rho)\rho}{1-\rho^{N+1}} \frac{d}{d\rho} \left(\frac{1-\rho^{N+1}}{1-\rho}\right) \\ &= \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})}, \quad \rho \neq 1 \end{aligned}$$

当 $\rho = 1$ 时, $L_s = \frac{N}{2}$(请验证!). 可以如 12.6.1 节那样, 从 L_s 利用 λ_{eff} 求出 W_s、W_q 和 L_q.

用计算器来计算排队公式相当麻烦 (后面的模型可能更麻烦!), 因此我们建议用 TORA 或电子表格程序 excelPoissonQ.xls 来解决计算问题.

例 12.6-4

考虑例 12.6-2 中的洗车设施. 假设该设施总共有 4 个停车位, 如果停车场满了, 新来到的车辆就会去别的洗车房. 老板想要确定有限停车位造成的顾客流失对竞争的影响.

按照模型的符号, 系统上限为 $N = 4 + 1 = 5$. 下面的输入数据提供了图 12.6 的输出.

Scenario 1:(M/M/1):(GD/5/infinity)

Lambda =	4.00000		Mu =	6.00000	
Lambda eff =	3.80752		Rho/c =	0.66667	
Ls =	1.42256		Lq =	0.78797	
Ws =	0.37362		Wq =	0.20695	
n	Probability pn	Cumulative Pn	n	Probability pn	Cumulative Pn
0	0.36541	0.36541	3	0.10827	0.87970
1	0.24361	0.60902	4	0.07218	0.95188
2	0.16241	0.77143	5	0.04812	1.00000

图 12.6　例 12.6-4 的 TORA 输出 (文件 toraEx12.6-4.txt)

Lambda	Mu	c	系统上限	顾客源上限
4	6	1	5	无限

因为系统上限为 $N = 5$, 流失顾客的比例为 $p_5 = 0.048\ 12$, 按照每天 24 小时算, 等价于每天损失 $(\lambda p_5) \times 24 = 4 \times 0.048\ 12 \times 24 = 4.62$ 辆车. 关于增加停车场规模的决策应根据损失业务的价值来作出.

从不同的角度来看这个问题, 系统中的期望总时间 W_s 为 0.373 6 小时, 大约为 22 分钟, 这比当允许所有到达车辆加入服务设施时的例 12.6-3 中的 30 分钟有所下降. 这一约 25% 的下降, 其代价是由于有限的停车位而流失了 4.8% 的潜在顾客.

习题 12.6D

***1.** 在例 12.6-4 中, 求下列各项:

(a) 到来的车辆马上能进入清洗位的概率.　(b) 服务开始前的期望等待时间.

(c) 期望空停车位数.　(d) 所有停车位都占满的概率.

(e) 限制系统中的平均时间为大约 10 分钟, 求平均服务时间降低百分率. (**提示**: 用 excelPoissonQ.xls 或 TORA 中的试错功能.)

2. 考虑例 12.6-4 中的洗车服务设施, 求所需要的停车位数, 使得找不到停车位的车辆比例不超过 1%.

3. 理发师 Joe 开的理发店为每位顾客理发的时间服从平均 12 分钟的指数分布. 由于他的理发店很有名, 来这儿的顾客 (服从泊松分布) 通常会比他的接待能力数多很多: 每小时 6 位顾客. 假如顾客到达率能有效地减少到每小时 4 位顾客的话, 他的理发店就能够应付自如了. 为了达到这一目的, Joe 提出了一个想法是, 在等待区提供有限个座位, 当新到来的顾客看到所有的座位都有人时, 他就会去其他地方了. 要实现 Joe 的目的, 他该提供多少个座位呢?

***4.** Electro 电气公司生产发电机的最后组装服从泊松分布, 每小时组装 10 台发电机. 然后把发电机放在传送带上送到质检部门做最后的调试. 传送带最多能放 7 台发电机. 一旦放满了, 有一个电子传感器会自动让传送带中断下来, 以防止最后组装线组装过量, 直到有空位置再启动. 发电机的检查时间服从指数分布, 平均检查时间为 15 分钟.

(a) 最后组装线停止生产的概率为多少? (b) 传送带上的发电机平均数是多少?

(c) 据生产部门的工程师称, 可以通过提高传送带能力来减少最终组装线的中断率. 事实上, 工程师的意义是, 传送带的能力可以增加到一定程度让组装线在 95% 的时候不发生中断. 他的话对吗?

5. 某自助餐厅最多可坐 50 人, 顾客按照泊松流到达, 到达率是每小时 10 个顾客, 并以每小时 12 位顾客的速度提供服务 (每次服务 1 人).

(a) 一位来到的顾客因为满员而不能在自助餐厅吃饭的概率是多少?

(b) 假设有 3 位顾客 (按随机到达时间) 想要坐在一起吃饭, 他们的愿望能得到满足的概率是多少?(假设只要有 3 个座位, 就可以把他们安排在一起)

6. 患者按照泊松分布来到一家有一个医生的诊所, 到达率是每小时 20 位患者. 候诊室最多能容纳 14 位患者. 每个患者的体检时间服从指数分布, 平均 8 分钟.

(a) 一位来到的患者不用等待的概率是多少?

(b) 一位来到的患者能在候诊室找到座位的概率是多少?

(c) 一位患者在该诊所花的期望总时间是多长?

7. 下表给出 $(M/M/1):(GD/5/\infty)$ 系统中有 n 个顾客的概率 p_n:

n	0	1	2	3	4	5
p_n	0.399	0.249	0.156	0.097	0.061	0.038

到达率 λ 为每小时 5 位顾客, 服务率 μ 是每小时 8 位顾客. 求:

*(a) 一来到的顾客能够进入系统的概率. *(b) 来到的顾客不能进入系统的比率.

(c) 系统中顾客的期望数. (d) 队列平均等待时间.

8. 证明对 $(M/M/1):(GD/N/\infty)$, 当 $\rho = 1$ 时, 系统中顾客的期望数 L_s 等于 $\frac{N}{2}$. [提示: $1+2+\cdots+i=\frac{i(i+1)}{2}$.]

9. 证明对 $(M/M/1):(GD/N/\infty)$, λ_{eff} 的计算公式为

$$\lambda_{\text{eff}} = \mu(L_s - L_q)$$

12.6.3 多服务台模型

本节考虑有多个并行服务台的 3 个排队模型. 前 2 个模型是 12.6.2 节模型的多服务台版本, 第 3 个模型针对自助服务的情况, 等价于有无限个并行的服务台.

实际应用——Qantas 航空公司电话售票人力资源计划问题

为了降低经营成本, Qantas 航空公司想要在本部电话订票办公室合理地安排工作人员, 以为顾客提供方便的服务. 按照传统方法, 对员工的需求的估计是根据过去的业务增长情况来预测未来电话呼叫量, 需要增加的员工数可用所预测的电话呼叫增长量除以一个接线员能够处理的平均呼叫数计算出来. 由于计算是基于平均数的, 增加雇用的员工数并没有考虑到每天需求的变化情况. 特别是, 在高峰业务时间内, 对服务的较长等待时间导致了顾客的抱怨, 并失去了一些业务. 要解决这

个问题, 就需要制定一个计划, 使得在所雇用的接线员数与顾客需求之间达到某种平衡. 解决这个问题采用了一个嵌入到整数规划模型中的 $(M/M/c)$ 排队分析. 仅在悉尼总部办公室用这一模型在 1975 到 1976 财政年度所节省的费用就达到 17.3 万美元. 这项研究的细节请见第 24 章的案例 15.

$(M/M/c)$:$(GD/\infty/\infty)$　在这个模型中有 c 个并行服务台. 到达率为 λ, 每个服务台的服务率为 μ. 因为对系统中的排队人数没有限制, 所以 $\lambda_{\text{eff}}=\lambda$.

使用并行服务台的效果是使得设施的服务率成比例增加. 根据广义模型 (12.5 节), λ_n 和 μ_n 定义为

$$\lambda_n=\lambda,\quad n\geqslant 0$$

$$\mu_n=\begin{cases} n\mu, & n<c \\ c\mu, & n\geqslant c \end{cases}$$

因此得到

$$p_n=\begin{cases} \dfrac{\lambda^n}{\mu(2\mu)(3\mu)\cdots(n\mu)}p_0=\dfrac{\lambda^n}{n!\mu^n}p_0=\dfrac{\rho^n}{n!}p_0, & n<c \\ \dfrac{\lambda^n}{\left(\prod\limits_{i=1}^{c} i\mu\right)(c\mu)^{n-c}}p_0=\dfrac{\lambda^n}{c!c^{n-c}\mu^n}p_0=\dfrac{\rho^n}{c!c^{n-c}}p_0, & n\geqslant c \end{cases}$$

令 $\rho=\frac{\lambda}{\mu}$, 并设 $\frac{\rho}{c}<1$, p_0 的值可从 $\sum\limits_{n=0}^{\infty}p_n=1$ 求出, 得到

$$\begin{aligned} p_0&=\left\{\sum_{n=0}^{c-1}\frac{\rho^n}{n!}+\frac{\rho^c}{c!}\sum_{n=c}^{\infty}\left(\frac{\rho}{c}\right)^{n-c}\right\}^{-1} \\ &=\left\{\sum_{n=0}^{c-1}\frac{\rho^n}{n!}+\frac{\rho^c}{c!}\left(\frac{1}{1-\frac{\rho}{c}}\right)\right\}^{-1},\quad \frac{\rho}{c}<1 \end{aligned}$$

L_q 的表达式如下式给出:

$$\begin{aligned} L_q&=\sum_{n=c}^{\infty}(n-c)p_n \\ &=\sum_{k=0}^{\infty}kp_{k+c} \\ &=\sum_{k=0}^{\infty}k\frac{\rho^{k+c}}{c^k c!}p_0 \\ &=\frac{\rho^{c+1}}{c!c}p_0\sum_{k=0}^{\infty}k\left(\frac{\rho}{c}\right)^{k-1} \\ &=\frac{\rho^{c+1}}{c!c}p_0\frac{d}{d\left(\frac{\rho}{c}\right)}\sum_{k=0}^{\infty}\left(\frac{\rho}{c}\right)^k \\ &=\frac{\rho^{c+1}}{(c-1)!(c-\rho)^2}p_0 \end{aligned}$$

由于 $\lambda_{\text{eff}} = \lambda$, $L_s = L_q + \rho$. W_s 和 W_q 的值可以用 L_s 和 L_q 除以 λ 而得到.

例 12.6-5

某社区由两家出租车公司提供服务, 每家公司有两辆出租车, 且这两家公司平等分享市场. 事实上, 叫车电话到达每家公司的派车办公室的到达率为每小时 8 次. 每次乘车的平均时间为 12 分钟. 叫车电话按照泊松分布到达, 乘车时间服从指数分布. 最近, 这两家公司被一个投资商购买了, 他打算把这两个派车办公室合成一个, 以便为顾客提供更优质的服务. 请分析这位新老板的建议.

从排队论的角度来看这个问题, 出租车就是服务台, 乘坐出租车就是服务. 每家公司都可以表示成参数为每小时 $\lambda = 8$ 次呼叫、每小时*每辆出租车* $\mu = \frac{60}{12} = 5$ 次乘坐的 $(M/M/2):(GD/\infty/\infty)$ 模型. 合并以后得到 $(M/M/4):(GD/\infty/\infty)$, 参数为每小时 $\lambda = 2 \times 8 = 16$ 次叫车, 每小时*每辆出租车* $\mu = 5$ 次乘坐.

比较这两个模型, 一种合适的度量指标是一个顾客等待乘车的平均时间 W_q. TORA 比较分析的输入数据如下表:

情景	Lambda	Mu	c	系统上限	输入源上限
1	8	5	2	无限	无限
2	16	5	4	无限	无限

图 12.7 给出了这两种情景下的输出结果. 这个结果说明, 等待乘车的时间在两台出租车情况下是 0.356 小时 (约等于 21 分钟), 而合并以后情况下为 0.149 小时 (约等于 9 分钟), 明显减少了 50% 多, 两个公司合并的效果非常明显.

Comparative analysis

c	Lambda	Mu	L'da eff	p0	Ls	Ws	Lq	Wq
2	8.000	5.000	8.00	0.110	4.444	0.556	2.844	0.356
4	16.000	5.000	16.00	0.027	5.586	0.349	2.386	0.149

图 12.7 例 12.6-5 的 TORA 输出 (文件 toraEx12.6-5.txt)

评注 以上分析的结论是, **共同分担服务***总是*一种更加有效的运作模式. 即使分别运营也"非常"忙的情况下, 这一结果也是对的 (见习题 12.6E 的第 2 题和第 10 题).

习题 12.6E

1. 考虑例 12.6-5.

(a) 说明对合并情况下等待时间明显减少 50% 以上是和提高服务繁忙时间的百分比密不可分的.

(b) 为了让一次乘车的平均等待时间限制在 5 分钟以下, 求合并后公司应该拥有的出租车辆数.

***2.** 在出租车公司的例子中, 假设每次乘车的平均时间大约在 14.5 分钟, 使得用 2 辆车和 4 辆车运营的利用率 ($= \frac{\lambda}{\mu c}$) 增加到 96% 以上, 还是否值得把两个公司合并成一个? 用一次乘

车的平均等待时间作为比较的度量指标.

3. 在下列 (泊松到达/离开) 每个情形下, 求最少的并行服务台数, 以保证排队系统的运行是稳定的 (即队列长度不会无限增长).

(a) 顾客每 5 分钟到达一位, 服务率为每小时 10 位顾客.

(b) 平均到达间隔时间是 2 分钟, 并且平均服务时间为 6 分钟.

(c) 到达率为每小时 30 位顾客, 每个服务台的服务率为每小时 40 位顾客.

4. 顾客按照泊松分布到达 Thrift 银行, 平均每小时 45 位顾客. 每个顾客的服务时间大约为 5 分钟, 服从指数分布. 银行想要采用一种单队列多营业窗口的运行方式, 类似于机场和邮局所采用的方式. 银行经理知道, 假如顾客感觉到他们的排队等待时间"过长"的话, 他们就可能会转到别的银行去了. 为此, 经理想要把排队的平均等待时间限制在不超过 30 秒钟. 银行应该提供多少营业窗口呢?

***5.** McBurger 快餐店有 3 个收款台. 顾客按照每 3 分钟/位的泊松分布到达, 排成一队等待第一个空闲的收款台提供服务. 点菜时间服从指数分布, 平均 5 分钟. 餐厅内的排队室空间有限, 但由于菜做得好, 必要时顾客们还是愿意在餐厅外排队. 求餐厅内排队室应设置多大, 使得顾客不用在餐厅外等待的概率至少达到 0.999.

6. 一家小邮局有 2 个开放窗口, 顾客按照泊松分布到达, 平均 3 分钟/位. 但是只有 80% 的顾客寻求邮局窗口服务. 每位顾客的服务时间服从指数分布, 平均 5 分钟. 所有来到的顾客排成一队, 按照先来先服务的规则在空出的窗口接受服务.

(a) 来到的顾客要排队等待的概率是多少?

(b) 2 个窗口都空闲的概率是多少?

(c) 等待队列的平均队长是多少?

(d) 能否只用一个窗口提供满意的服务? 请解释.

7. A 大学计算中心配置了 4 台相同型号的大型机. 任何时候的用户数都是 25 人. 任何一个用户都能够平均 15 分钟从终端提交一份作业, 但提交作业之间的实际时间服从指数分布, 提交来的作业自动进入到第一台空闲的计算机, 每份作业的执行时间也服从指数分布, 平均 2 分钟. 求下列各项:

*(a) 一份作业提交后不能马上执行的概率.

(b) 一份作业的输出返回到用户的平均时间.

(c) 等待执行的平均作业数.

(d) 整个计算中心都空闲的时间百分比.

*(e) 空闲计算机的平均数.

8. Drake 机场为乡村、郊区和转机的乘客提供服务. 这 3 类旅客的到达时间服从泊松分布, 平均每小时分别为 15, 10, 20 位旅客. 每位旅客办理登机手续的时间服从指数分布, 平均 6 分钟. 求 Drake 机场应提供的满足下列一个条件的柜台数.

(a) 每位旅客办理乘机手续的总时间小于 15 分钟.

(b) 柜台的空闲率不超过 10%.

(c) 所有柜台都空闲的概率不超过 0.01.

9. 在美国, 邮局和机场办理登机延续的柜台常常采用单队列多服务台的排队方式. 但是, (特别是小社区的) 杂货店和银行更愿意采用单队列和单服务台的方式, 虽然单队列多服务台排队能提供更高效的操作. 对这一现象给出评价.

10. 对 $(M/M/c):(GD/\infty/\infty)$ 模型, 莫尔斯 (1958, p.103) 证明, 当 $\frac{\rho}{c}\to 1$ 时,

$$L_q = \frac{\rho}{c-\rho}$$

注意 $\frac{\rho}{c}\to 1$ 意味着服务台非常繁忙, 利用这一信息说明, 当 $\frac{\rho}{c}\to 1$ 时, $(M/M/c):(GD/\infty/\infty)$ 模型的平均队列等待时间与 $(M/M/1):(GD/\infty/\infty)$ 模型的平均队列等待时间之比趋向于 $\frac{1}{c}$. 因此对于 $c=2$, 平均等待时间可以减少 50%. 从本题得出的结论是, 不论服务台如何“超负荷工作”, 合并服务永远是更好的.

11. 在得出 $(M/M/c):(GD/\infty/\infty)$ 模型的 p_n 时, 说明哪一步骤需要条件 $\frac{\rho}{c}<1$. 解释这个条件的含义. 如果条件不满足将会怎样?

12. 从定义 $L_q=\sum\limits_{n=c+1}^{\infty}(n-c)p_n$ 出发, 证明 $L_s=L_q+\bar{c}$, 其中 $\bar{c}$ 为繁忙服务台的平均数. 因此说明 $\bar{c}=\frac{\lambda_{\text{eff}}}{\mu}$.

13. 说明对于 $(M/M/1):(GD/\infty/\infty)$ 模型, p_n 可以从 $(M/M/c):(GD/\infty/\infty)$ 模型通过设定 $c=1$ 得出.

14. 对于 $(M/M/c):(GD/\infty/\infty)$ 模型, 证明

$$L_q=\frac{c\rho}{(c-\rho)^2}p_c$$

15. 对于 $(M/M/c):(GD/\infty/\infty)$ 模型, 证明

(a) 顾客等待的概率为 $\frac{\rho}{(c-\rho)}p_c$.

(b) 已知非空队列中的顾客平均数为 $\frac{c}{(c-\rho)}$.

(c) 对必须等待的顾客的期望队列等待时间是 $\frac{1}{\mu(c-\rho)}$.

16. 证明 $(M/M/c):(GD/\infty/\infty)$ 模型队列等待时间的概率密度函数为

$$w_q(T)=\begin{cases}1-\dfrac{\rho^c}{(c-1)!(c-\rho)}p_0, & T=0\\ \dfrac{\mu\rho^c\mathrm{e}^{-\mu(c-\rho)T}}{(c-1)!}p_0, & T>0\end{cases}$$

[提示: 把 c 个通道情形转换成等价的单通道, 利用

$$\mathrm{P}\{t>T\}=\mathrm{P}\{\min_{1\leqslant i\leqslant c}t_i>T\}=(\mathrm{e}^{-\mu T})^c\mathrm{e}^{-\mu cT}$$

其中 t 是等价的单通道中的等待时间].

17. 对第 16 题中的 $w_q(T)$, 证明

$$\mathrm{P}\{T>y\}=\mathrm{P}\{T>0\}\mathrm{e}^{-(c\mu-\lambda)y}$$

其中 $\mathrm{P}\{T>0\}$ 为来到的顾客必须等待的概率.

18. 证明对 $(M/M/c):(FCFS/\infty/\infty)$ 模型的系统等待时间有如下概率密度函数:

$$w(\tau)=\mu\mathrm{e}^{-\mu\tau}+\frac{\rho^c\mu\mathrm{e}^{-\mu\tau}}{(c-1)!(c-\rho-1)}\left\{\frac{1}{c-\rho}-\mathrm{e}^{-\mu(c-\rho-1)\tau}\right\}p_0,\quad \tau\geqslant 0$$

(提示: τ 是队列中等待时间 T 和服务时间分布的卷积.)

$(M/M/c):(GD/N/\infty), c\leqslant N$ 这个模型与 $(M/M/c):(GD/\infty/\infty)$ 模型的不同在于, 系统上限是有限的并且等于 N. 这就意味着, 最大队列长度是 $N-c$. 到达率和服务率分别为 λ 和 μ. 因为系统上限为 N, 因此有效到达率 λ_{eff} 小于 λ.

按照广义模型 (12.5 节), 当前模型的 λ_n 和 μ_n 定义为

$$\lambda_n = \begin{cases} \lambda, & 0\leqslant n\leqslant N \\ 0, & n>N \end{cases}, \quad \mu_n = \begin{cases} n\mu, & 0\leqslant n\leqslant c \\ c\mu, & c\leqslant n\leqslant N \end{cases}$$

在 12.5 节的一般表达式中替代 λ_n 和 μ_n, 并注意到 $\rho=\frac{\lambda}{\mu}$, 得到

$$p_n = \begin{cases} \dfrac{\rho^n}{n!}p_0, & 0\leqslant n< c \\ \dfrac{\rho^n}{c!c^{n-c}}p_0, & c\leqslant n\leqslant N \end{cases}$$

其中

$$p_0 = \begin{cases} \left(\displaystyle\sum_{n=0}^{c-1}\frac{\rho^n}{n!} + \frac{\rho^c\left(1-\left(\frac{\rho}{c}\right)^{N-c+1}\right)}{c!(1-\frac{\rho}{c})}\right)^{-1}, & \dfrac{\rho}{c}\neq 1 \\ \left(\displaystyle\sum_{n=0}^{c-1}\frac{\rho^n}{n!} + \frac{\rho^c}{c!}(N-c+1)\right)^{-1}, & \dfrac{\rho}{c}=1 \end{cases}$$

接下来, 计算 $\frac{\rho}{c}\neq 1$ 情况下的 L_q:

$$\begin{aligned} L_q &= \sum_{n=c}^{N}(n-c)p_n \\ &= \sum_{j=0}^{N-c} jp_{j+c} \\ &= \frac{\rho^c\rho}{c!c}p_0\sum_{j=0}^{N-c} j\left(\frac{\rho}{c}\right)^{j-1} \\ &= \frac{\rho^{c+1}}{cc!}p_0\frac{d}{d\left(\frac{\rho}{c}\right)}\sum_{j=0}^{N-c}\left(\frac{\rho}{c}\right)^j \\ &= \frac{\rho^{c+1}}{(c-1)!(c-\rho)^2}\left\{1-\left(\frac{\rho}{c}\right)^{N-c+1}-(N-c+1)\left(1-\frac{\rho}{c}\right)\left(\frac{\rho}{c}\right)^{N-c}\right\}p_0 \end{aligned}$$

可以证明当 $\frac{\rho}{c}=1$ 时, L_q 化简为

$$L_q = \frac{\rho^c(N-c)(N-c+1)}{2c!}p_0, \quad \frac{\rho}{c}=1$$

为了求 W_q, 进而求出 W_s 和 L_s, 所计算的 λ_{eff} 的值为

$$\lambda_{\text{lost}} = \lambda p_N$$

$$\lambda_{\text{eff}} = \lambda-\lambda_{\text{lost}} = (1-p_N)\lambda$$

例 12.6-6

在例 12.6-5 的合并出租车公司问题中, 假设没有新的经费用来购买更多出租车. 有一个咨询专家向老板建议, 有一种减少等待时间的方法是, 一旦有 6 个顾客在等待用车了, 就让派车办公室通知新的顾客, 告诉他们等待时间可能会很长. 这种举动定能会让新的顾客去寻求别的公司的服务, 但将会减少等待顾客的等待时间. 请评价这位专家的建议.

将等待的顾客数限制在 6 个以内就等价于设定 $N = 6 + 4 = 10$ 个顾客. 因此我们来考察模型 $(M/M/4) : (GD/10/\infty)$, 其中 $\lambda = 16$ 个顾客/小时, $\mu = 5$ 次乘车/小时. 从下面的输入数据可以得出图 12.8 的结果.

Lambda	Mu	c	系统上限	顾客源上限
16	5	4	10	无限

```
Scenario1:(M/M/4):(GD/10/infinity)
Lambda =        16.00000          Mu =        5.00000
Lambda eff =    15.42815          Rho/c =     0.80000
Ls =             4.23984          Lq =        1.15421
Ws =             0.27481          Wq =        0.07481
```

n	Probability pn	Cumulative Pn	n	Probability pn	Cumulative Pn
0	0.03121	0.03121	6	0.08726	0.79393
1	0.09986	0.13106	7	0.06981	0.86374
2	0.15977	0.29084	8	0.05584	0.91958
3	0.17043	0.46126	9	0.04468	0.96426
4	0.13634	0.59760	10	0.03574	1.00000

图 12.8 例 12.6-6 的 TORA 输出 (文件 toraEx12.6-6.txt)

在设置系统能力上限之前的平均等待时间 W_q 为 0.149 小时 (≈ 9 分钟) (见图 12.7), 大约是新的平均等待时间 0.075 小时 (≈ 4.5 分钟) 的两倍. 等待时间的这一大幅度减少的代价是流失了大约 3.6% 的潜在顾客 ($p_{10} = 0.035\,74$). 但是这个结果还不能反映顾客对公司经营印象的损害效果.

习题 12.6F

1. 在例 12.6-6 中, 求下列各项:

   (a) 空闲出租车的期望数. (b) 某叫车顾客是排队顾客最后一个的概率.

   (c) 假如要将队列的等待时间限制在 3 分钟以下的话, 求对于等待队列长度的限制.

2. Eat&Gas 便利店经营一个有 2 个油泵的加油站, 不包括正在加油的车辆, 通向加油泵的车道最多可供 3 辆车驶入. 如果车道占满了, 来到的汽车就要去别的地方了. 来到的车辆服从泊松分布, 平均每小时 20 辆车. 加满油并付费的时间服从指数分布, 平均 6 分钟. 求下列各项:

   (a) 到别处寻找服务的车辆的百分比. (b) 1 台加油泵在使用的时间百分比.

*(c) 2 台油泵的利用率百分比.

*(d) 来到的车辆没有立刻开始接受服务, 但找到了空车道位置的概率.

(e) 求车道容量, 使得保证平均不超过 10% 的到达车辆走开.

(f) 求车道容量, 使得 2 台油泵都空闲的概率小于等于 0.05%.

3. 3 个机械工合开一家小型的机器修理店. 每年 3 月份, 人们都把他们的犁耙和除草机拿来维修. 修理厂愿意接受顾客拿来的所有犁耙和除草机, 但每当新的顾客看到厂房堆满了等待维修的农具, 他们就到别的能得到较快修理的地方去了. 不包括正在修理的机器, 维修间最多能放 15 台除草机或犁耙. 平均每 10 分钟有一位顾客来到修理厂, 每个机械工平均要花 30 分钟修好一件农具. 顾客到达间隔时间和服务时间都服从指数分布, 求下列各项:

(a) 空闲机械工的平均数.

(b) 由于店里工作时间有限, 每天按 10 小时算, 流失的顾客量有多少?

(c) 下一个来到的顾客能被修理店服务的概率.

(d) 至少有一个机械工空闲的概率.

(e) 等待修理的犁耙或除草机的平均数. (f) 该修理店的总体生产率度量.

4. A 大学新生总想开车去上课 (尽管学校要求大部分学生住在校园内, 可以利用学校的免费交通工具). 在秋季学期的前几周里, 由于新生们到处寻找停车位, 造成校园内交通十分拥堵. 尽管如此, 学生们在停车场的车道内仍耐心等待, 等候有人离开, 好把自己的车子停进去. 让我们来想象一个情景: 停车场有 30 个停车位, 车道上还可以停 10 辆车, 但这 10 辆车不能永远停放在车道上, 必须等待这 30 个车位有人倒出来. 这些新生开车按照泊松分布来到停车场, 平均每小时 20 辆, 每辆车的平均泊车时间大约是 60 分钟, 但实际时间服从指数分布.

*(a) 由于进不去停车场而走开的新生比例是多少?

*(b) 一辆来到的车必须在车道内等待的概率是多少?

(c) 一辆来到的车占据停车场仅有的一个车位的概率是多少?

*(d) 求占据的停车位的平均数.

(e) 求车道内被占据的车位的平均数.

*(f) 求由于停车场全满了而在 8 小时期间没上成课的新生数.

5. 验证 $(M/M/c):(GD/N/\infty)$ 模型当 $\frac{\rho}{c}\neq 1$ 时 p_0 的公式.

6. 对 $(M/M/c):(GD/N/\infty)$ 模型, 证明下列等式

$$\lambda_{\text{eff}} = \mu\bar{c}$$

其中 $\bar{c}$ 为繁忙服务台数.

7. 验证 $(M/M/c):(GD/N/\infty)$ 模型当 $\frac{\rho}{c}=1$ 时 p_0 和 L_q 的公式.

8. 对 $(M/M/c):(GD/N/\infty)$ 模型, 其中 $N=c$, λ_n, μ_n 的定义同广义模型 (12.5 节), 证明 p_n 的公式为

$$p_n = \frac{\rho^n}{n!}p_0, \quad n=1,2,\cdots,c$$

其中

$$p_0 = \left(1+\sum_{n=1}^{c}\frac{\rho^n}{n!}\right)^{-1}$$

$(M/M/\infty)$:$(GD/\infty/\infty)$——自助服务模型 在这个模型中, 因为顾客本身也是服务台, 因此服务台数量无限. 该模型的一个典型例子是参加驾驶执照考试的笔试部分. 自助式加油站和 24 小时自动取款机银行服务都不符合这个模型的描述, 因为这两个情况中, 实际的服务台是加油泵和 ATM 机. 这个模型假定稳定的到达率和服务率, 分别为 λ 和 μ.

按照 12.5 节的广义模型, 有

$$\lambda_n = \lambda, \quad n = 0, 1, 2, \cdots$$
$$\mu_n = n\mu, \quad n = 0, 1, 2, \cdots$$

因此有

$$p_n = \frac{\lambda^n}{n!\mu^n} p_0 = \frac{\rho^n}{n!} p_0, \quad n = 0, 1, 2, \cdots$$

由于 $\sum\limits_{n=0}^{\infty} p_n = 1$, 得到

$$p_0 = \frac{1}{1 + \rho + \frac{\rho^2}{2!} + \cdots} = \frac{1}{e^{\rho}} = e^{-\rho}$$

结果有

$$p_n = \frac{e^{-\rho}\rho^n}{n!}, \quad n = 0, 1, 2, \cdots$$

这是平均值为 $L_s = \rho$ 的泊松分布. 如我们期望的那样, 因为是自助服务模型, L_q 和 W_q 都等于零.

例 12.6-7

一个投资者每月平均投入 \$1 000 购买一种股票市场的债券. 因为这个投资者必须要等待好的“买入”机会, 实际发生的购买时间是完全随机的. 该投资者平均要把债券保留 3 年, 但是当好的“卖出”机会来了, 他会在随机的时间把它卖掉. 尽管这个投资者一般被认为是一个精明的股票市场玩家, 但过去的经验表明, 大约有 25% 的债券每年下跌约 20% 左右, 其余的 75% 每年上涨 12% 左右. 请估算一下这位投资者在股票市场的 (长期) 平均资产净值.

这一情形可以看成是一个 $(M/M/\infty):(GD/\infty/\infty)$ 模型, 因为从实际情况考虑, 这位投资者并不需要排队等待债券的买入或卖出. 买卖间隔的平均时间是 1 个月, 因此每年有 $\lambda = 12$ 只债券. 债券的销售率为每年 $\mu = \frac{1}{3}$ 只债券. 你可以用下面的输入来得到模型的结果:

Lambda	Mu	c	系统上限	顾客源上限
12	0.333 333 3	无限	无限	无限

已知 λ 和 μ 的值, 得到

$$L_s = \rho = \frac{\lambda}{\mu} = 36 \text{ 只债券}$$

该投资者的预计 (长期) 平均年度净值为

$$(0.25L_s \times \$1\,000)(1-0.20) + (0.75L_s \times \$1\,000)(1+0.12) = \$63\,990$$

习题 12.6G

1. 在例 12.6-7 中, 计算下列各项:

(a) 该投资者完全卖空的概率.　**(b)** 该投资者拥有至少 10 只债券的概率.

(c) 该投资者会拥有 30 到 40 只债券的概率.

(d) 假如只有 10% 的债券贬值 30%, 其余的 90% 债券每年升值 15%, 求该投资者的年度债券净值.

2. 在道路驾驶考试之前, 要求所有的新司机都要通过笔试. 这些考试通常由市警察局主管. Springdale 市的记录显示, 每天 8 小时内来参加笔试的平均人数有 100 人, 完成考试所需要的平均时间为 30 分钟. 但是实际来参加考试的人数和每个人所花费的考试时间是完全随机的. 求下列各项:

***(a)** 警察局应在考试大厅设置座位的平均数.

***(b)** 参加考试的人数超过考试大厅座位平均数的概率.

(c) 任何一天没有人来参加考试的概率.

3. (利用 excelPoissonQ.xls 或 TORA) 说明对于很小的 $\rho = 0.1$, $(M/M/c):(GD/\infty/\infty)$ 模型中的 L_s, L_q, W_s, W_q, p_n 的值, 对于 c 小于等于 4 个服务台时, 可以通过 $(M/M/\infty):(GD/\infty/\infty)$ 模型的简化公式可靠地估计出来.

4. 对于较大的 $\rho = 9$, 重复第 3 题并说明, 除了 c 的值必须更大以外 (至少 14), 结论同样成立. 从第 3 题和第 4 题的结果, 关于用 $(M/M/\infty):(GD/\infty/\infty)$ 模型来估计 $(M/M/c):(GD/\infty/\infty)$ 模型的结果, 可以得出什么一般性的结论?

12.6.4　机器侍服模型——$(M/M/R):(GD/K/K), R<K$

这个模型的背景是有 K 台机器的车间. 当一台机器出现故障时, 就呼叫 R 个有时间的修理工之一来进行修理. 每台机器的故障率为每单位时间 λ 次故障, 每个修理工修理故障机器的服务率为每单位时间 μ 台机器. 所有的故障和服务假定都服从泊松分布.

这个模型和前面介绍的模型都有所不同, 因为它有有限的输入源. 要看到这一点, 只要我们了解到当车间里所有的机器都坏了, 就不再叫修理工来服务了. 其实, 只有在工作的机器才会出故障, 因此才能够产生呼叫服务.

设每台机器的故障率为 λ, 则整个车间的故障率与正在工作的机器数成比例. 按照排队模型的术语, 系统中有 n 台机器就是说有 n 台机器都出现了故障, 因此整个车间的故障率就是

$$\lambda_n = (K-n)\lambda, \quad 0 \leqslant n \leqslant K$$

根据 12.5 节的广义模型, 有

$$\lambda_n = \begin{cases} (K-n)\lambda, & 0 \leqslant n \leqslant K \\ 0, & n \geqslant K \end{cases}$$

$$\mu_n = \begin{cases} n\mu, & 0 \leqslant n \leqslant R \\ R\mu, & R \leqslant n \leqslant K \end{cases}$$

从广义模型, 就可以得到 (请验证!)

$$p_n = \begin{cases} C_n^K \rho^n p_0, & 0 \leqslant n \leqslant R \\ C_n^K \dfrac{n!\rho^n}{R!R^{n-R}} p_0, & R \leqslant n \leqslant K \end{cases}$$

$$p_0 = \left(\sum_{n=0}^{R} C_n^K \rho^n + \sum_{n=R+1}^{K} C_n^K \frac{n!\rho^n}{R!R^{n-R}} \right)^{-1}$$

L_s 不存在闭型的表达式, 因此必须用下面的基本定义来计算:

$$L_s = \sum_{n=0}^{K} np_n$$

λ_{eff} 值的计算为

$$\lambda_{\text{eff}} = \mathrm{E}\{\lambda(K-n)\} = \lambda(K-L_s)$$

利用 12.6.1 节的公式, 可以算出其余的系统性能度量指标 W_s, W_q, L_q.

例 12.6-8

Toolco 公司经营一家有 22 台机床的机械加工厂. 已知每台机床平均每 2 个小时发生一次故障, 修理工作平均需要 12 分钟. 故障间隔时间和修理时间均服从指数分布. Toolco 想要确定需要的修理工数, 以保证工厂能够"平稳地"运转.

要分析这种情况, 考察作为修理工数的函数的机床生产率, 这一生产率的度量可以定义为

$$\text{机床生产率} = \frac{\text{所有的机床} - \text{发生了故障的机床}}{\text{所有的机床}} \times 100$$

$$= \frac{22 - L_s}{22} \times 100$$

用下列输入数据, 可得出对该情形的分析结果: lambda= 0.5, mu= 5, $R = 1, 2, 3, 4$, 系统上限 = 22, 输入源 = 22. 图 12.9 给出了输出结果. 相应的生产率计算如下:

```
Comparative Analysis
c    Lambda    Mu      L'da eff      p0         Ls         Lq         Ws         Wq
1    0.500     5.00    4.9980     0.0004    12.0040    11.0044     2.4018     2.2018
2    0.500     5.00    8.8161     0.0564     4.3677     2.6045     0.4954     0.2954
3    0.500     5.00    9.7670     0.1078     2.4660     0.5128     0.2525     0.0525
4    0.500     5.00    9.9500     0.1199     2.1001     0.1102     0.2111     0.0111
```

图 12.9 例 12.6-8 的 TORA 比较分析输出 (文件 toraEx12.6-8.txt)

修理工 R	1	2	3	4
机床生产率 (100%)	45.44	80.15	88.79	90.45
边际增长率 (100%)	—	34.71	8.64	1.66

这些结果说明, 用 1 个修理工时, 生产率很低 (= 45.44%). 把修理工数增加到 2 个时, 生产率增加到 80.15%, 上升了 34.71%. 当我们雇用 3 个修理工时, 生产率只增加了 8.64%, 提高到 88.79%, 而 4 个修理工只把生产率增加一个很小量 1.66%, 提高到 90.45%.

从这些结果可以判断出, 用 2 个修理工最划算, 用 3 个修理工的效果不明显, 因为生产率仅仅提高了 8.64%. 当然我们可以用经费上的比较来确定是否合算, 这需要对雇用第 3 个修理工的成本和这 8.64% 的生产率提高所带来的收入进行比较 (见 12.9.1节对费用模型的讨论). 至于雇用第 4 个修理工, 生产率的微量增长 1.66% 不支持这样的行动.

习题 12.6H

1. 在例 12.6-8 中, 进行下列计算:

(a) 根据图 12.9, 验证 λ_{eff} 的值. *(b) 已知 R=4, 计算空闲修理工的期望数.

(c) 已知 R=3, 计算所有修理工都空闲的概率.

*(d) 已知 R=3, 计算大多数 (一半以上) 修理工空闲的概率.

2. 在例 12.6-8 中, 定义和计算 $R = 1, 2, 3, 4$ 时修理工的生产率. 利用这些信息连同机床生产率, 决定 Toolco 公司应该雇用的修理工数.

3. 在图 12.9 的计算中, 工厂中机床的平均故障率 λ_{eff} 随着 R 的增加而增加, 这一点可能不太清楚. 请解释为什么 λ_{eff} 会这样增加呢?

***4.** 一个操作工人照看 5 台自动机床. 在一台机床完成一批加工作业之后, 这位操作工必须要在新一批加工作业开始前重新设置机床. 完成一批加工作业的时间服从指数分布, 平均 45 分钟. 机床设置时间也服从指数分布, 平均 8 分钟.

(a) 求等待重新设置或正在设置的机床的平均数.

(b) 求所有机床都正在加工作业的概率.

(c) 求一台机床停止工作的平均时间.

5. Kleen All 是一家承接各种杂活的服务公司, 比如收拾庭院、树木剪枝、粉刷房屋等. 这家公司有 4 名工作人员, 每天接到工作后就离开办公室出去, 每人等工作干完了再给办公室打电话, 询问下一项工作的指令. 完成一次工作的时间服从指数分布, 平均 45 分钟. 两项相继工作之间的路上时间也服从指数分布, 平均 20 分钟.

(a) 求两项相继工作之间在路上的员工的平均数.

(b) 计算没有员工在路上的概率.

*6. 等待了很长时间以后, Newborn 夫妇终于迎来了五胞胎的出生: 两个男孩, 三个女孩. 在头 5 个月里, 这些宝宝的生活有两种状态组成: 醒着 (大多时间在哭) 和睡觉. 根据 Newborn 夫妇称, 孩子们的"醒着–睡着"活动从来不一致, 而是完全随机的. 事实上, 做统计工作的 Newborn 太太相信, 每个婴儿啼哭的时间服从指数分布, 平均是 30 分钟, 每个孩子的睡眠量也服从指数分布, 平均 12 个小时. 求下列各项:

(a) 求任何时候醒着的孩子的平均数. (b) 求所有该子都在睡觉的概率.

(c) 求由于醒着的孩子 (并且啼哭) 比睡觉的孩子多而让 Newborn 夫妇很郁闷的概率.

7. 对 $(M/M/R):(GD/K/K)$ 模型, 检验 p_n 的表达式.

8. 说明工厂的机床故障率可以按下面的公式计算:

$$\lambda_{\text{eff}} = \mu \bar{R}$$

其中 $\bar{R}$ 为繁忙修理工的平均数.

9. 针对一个修理工的特殊情况 $(R=1)$, 验证下列结果:

$$p_n = \frac{K!\rho^n}{(K-n)!}p_0$$

$$p_0 = \left(1 + \sum_{n=1}^{R} \frac{K!\rho^n}{(K-n)}\right)^{-1}$$

$$L_s = K - \frac{(1-p_0)}{\rho}$$

12.7 $(M/G/1)$:$(GD/\infty/\infty)$——Pollaczek-Khintchine(P-K) 公式

到达和离开都不服从泊松分布的排队模型很复杂. 一般情况下, 对这类情形的分析, 我们建议采用模拟的方法 (见第 16 章).

本节介绍一个很少出现的具有解析结果的非泊松排队. 它所描述的情况是, 服务时间 t 服从任何概率分布, 平均值为 $\mathrm{E}\{t\}$, 方差为 $\mathrm{Var}\{t\}$. 这个模型的结果包括系统性能的基本度量指标 L_s、L_q、W_s 和 W_q. 由于公式非常复杂, 这个模型没有 p_n 的闭型表达式.

令 λ 为单服务台设施的到达率, 已知服务时间分布的 $\mathrm{E}\{t\}$ 和 $\mathrm{Var}\{t\}$, 并且 $\lambda\mathrm{E}\{t\} < 1$, 我们可以通过复杂的概率论/马尔可夫链分析来证明

$$L_s = \lambda\mathrm{E}\{t\} + \frac{\lambda^2(\mathrm{E}^2\{t\} + \mathrm{Var}\{t\})}{2(1-\lambda\mathrm{E}\{t\})}, \quad \lambda\mathrm{E}\{t\} < 1$$

服务设施为空闲的概率可按下面公式求出:

$$p_0 = 1 - \lambda\mathrm{E}\{t\} = 1 - \rho$$

由 $\lambda_{\text{eff}} = \lambda$, 其他的系统性能度量指标 (L_q, W_s, W_q) 可以从 L_s 求出来, 如 12.6.1 节所解释的那样.

利用 excelPKFormula.xls 可以自动计算这个模型.

例 12.7-1

在例 12.6-2 的 Automata 洗车房问题中, 我们假设他们安装了一套新的洗车系统, 对所有车辆的服务时间均为一个常数, 等于 10 分钟. 那么这套新系统会对洗车房的运行有什么影响呢?

根据例 12.6-2, $\lambda_{\text{eff}} = \lambda = 4$ 辆车/小时. 服务时间为常数, 这样 $\text{E}\{t\} = \frac{10}{60} = \frac{1}{6}$ 小时, $\text{Var}\{t\} = 0$. 因此,

$$L_s = 4\left(\frac{1}{6}\right) + \frac{4^2\left[\left(\frac{1}{6}\right)^2 + 0\right]}{2\left(1 - \frac{4}{6}\right)} = 1.33 \text{ 辆车}$$

$$L_q = 1.333 - \left(\frac{4}{6}\right) = 0.667 \text{ 辆车}$$

$$W_s = \frac{1.333}{4} = 0.333 \text{ 小时}$$

$$W_q = \frac{0.667}{4} = 0.167 \text{ 小时}$$

有意思的是, 虽然到达率和离开率与例12.6-2 的泊松分布情况相同 ($\lambda = 4$ 辆车/小时, $\mu = \frac{1}{\text{E}\{t\}} = 6$ 辆车/小时), 但这个模型的期望等待时间要少, 因为服务时间为常数, 如下表所示.

	$(M/M/1):(GD/\infty/\infty)$	$(M/D/1):(GD/\infty/\infty)$
W_s(hr)	0.500	0.333
W_q(hr)	0.333	0.167

因为常数服务时间表示该设施的运行*更加确定*, 所以这些结果是合理的.

习题 12.7A

1. 在例 12.7-1 中, 计算设施空闲时间的百分比.

2. 求解例 12.7-1, 假定服务时间分布如下:

*(a) 8 至 20 分钟之间的均匀分布. **(b)** 正态分布, $\mu = 12$ 分钟, $\sigma = 3$ 分钟.

(c) 离散分布, 当值为 4, 8, 15 分钟时, 概率值分别为 0.2, 0.6, 0.2.

3. Layson 屋顶安装公司的业务是为阿肯色州的新老居民安装屋顶. 客户们随机地向该公司申请服务, 平均每月 30 天里有 9 家业务, 按照先来先服务的规则等待上门服务. 每家的面积不同, 但我们大致可以假定, 屋顶的面积服从 150 平方米到 300 平方米的均匀分布. 施工队一天能完成 75 平方米, 求下列各项:

(a) Layson 公司安装屋顶业务的平均积压作业数.

(b) 一个顾客直到安装完屋顶的平均等待时间.

(c) 假如增加施工队规模, 使得每天能完成 150 平方米, 这对完成作业的平均时间有什么影响?

***4.** Optica 有限公司根据顾客的订单配制处方眼镜. 每个工人只会配制某几种眼镜. 该公司配制的双焦距和三焦距眼镜积压了很多订单. 负责配制这种镜片的工人每天 8 小时会收到 30 份订单. 配制一副眼镜的时间服从正态分布, 平均值为 12 分钟, 标准差为 3 分钟. 在花 2 到 4 分钟 (均匀分布) 检查这些镜片以后, 该工人才能开始制作新的镜片. 求下列各项:

(a) 工人空闲时间的比例.

(b) Optica 公司双焦距和三焦距镜片的平均积压数.

(c) 完成一副镜片的平均时间.

5. 某产品按照泊松分布到达, 到达率为每 45 分钟一件. 该产品要求一名工人按两道工序加工. 第一道工序采用一台半自动机床加工, 恰好 28 分钟完成加工周期. 第二道工序只是作调整和少量的改进, 其加工时间的长短依赖于第一道工序加工后的产品条件. 特别地, 第二道工序加工时间服从 3 到 6 分钟的均匀分布. 由于每一道工序的加工都要求工人全神贯注, 必须要把第二道工序加工的工件卸下后, 才能在半自动机床上装上新的工件.

(a) 求在半自动机床上等待加工的工件数.

(b) 工人空闲时间的百分比例有多大.

(c) 一个来到的工件完成第二道工序的平均时间多长?

6. $(M/D/1):(GD/\infty/\infty)$. 说明在服务时间为常数的情况下, P-K 公式简化为

$$L_s = \rho + \frac{\rho^2}{2(1-\rho)}$$

其中 $\mu = \frac{1}{\mathrm{E}\{t\}}$ 且 $\rho = \frac{\lambda}{\mu} = \lambda \mathrm{E}\{t\}$.

7. $(M/E_m/1):(GD/\infty/\infty)$. 已知服务时间服从参数为 m 和 μ 的埃尔朗分布 (即, $\mathrm{E}\{t\} = \frac{m}{\mu}, \mathrm{Var}\{t\} = \frac{m}{\mu^2}$), 证明 P-K 公式化简为

$$L_s = m\rho + \frac{m(1+m)\rho^2}{2(1-m\rho)}$$

8. 当服务时间服从平均值为 $\frac{1}{\mu}$ 时间单位的指数分布时, P-K 公式化简为 $(M/M/1):(GD/\infty/\infty)$ 的 L_s.

9. 在一个有 c 个并行服务台的服务设施中, 假设顾客的到达服从泊松分布, 平均到达率为 λ. 来到的顾客按照一种严格的轮流方式指定给这些服务台 (不论繁忙还是空闲).

(a) 求顾客到达间隔时间的概率分布.

(b) 在 (a) 中进一步假设, 来到的顾客按照概率 $\alpha_i, \alpha_i \geqslant 0, i = 1, 2, \cdots, c$, 并且 $\alpha_1 + \alpha_2 + \cdots + \alpha_c = 1$, 指派给这 c 个服务台. 求顾客到达间隔时间的概率分布.

12.8 其他排队模型

虽然前面几节集中介绍了泊松排队模型, 排队系统还有很多其他类型的模型. 特别地, 带有服务优先权、网络队列以及非泊松队列的排队模型构成了排队论文献的重要组成部分. 这些模型出现在最专业的排队论著作中.

12.9 排队决策模型

排队设施的服务水平是服务率 μ 和并行服务台数 c 的函数. 本节介绍两个用来确定排队系统“合适的”服务水平的决策模型: (1) 费用模型; (2) 渴望水平模型. 这两个模型都认为, 提高服务水平会减少系统的等待时间. 而这两个模型的目的也都是在服务水平和等待的各个冲突因素之间寻求某种平衡.

12.9.1 费用模型

费用模型试图平衡两种冲突的费用:

(1) 提供服务的费用;

(2) 在提供服务中延误时间 (顾客等待时间) 的费用.

这两种费用是相互冲突的, 因为一个费用的增加, 必然引起另一个费用的减少, 如前面的图 12.1 所示.

令 $x(=\mu或c)$ 表示服务水平, 则费用模型可表示为

$$ETC(x)=EOC(x)+EWC(x)$$

其中

$$ETC=\text{单位时间的期望总费用}$$
$$EOC=\text{单位时间运行设施的期望费用}$$
$$EWC=\text{单位时间等待的期望费用}$$

EOC 和 EWC 的最简单形式是下面的线性函数:

$$EOC(x)=C_1x$$
$$EWC(x)=C_2L_s$$

其中

$C_1=$ 每单位时间每单位 x 的边际成本

$C_2=$ 每一位 (等待) 顾客每单位时间的等待费用

下面两个例子说明如何应用费用模型. 第 1 个例子假定 x 等于服务率 μ, 第 2 个例子假定 x 等于并行服务台数 c.

例 12.9-1

KeenCo 印刷公司正在采购一种高速商用印刷机. 有 4 个型号的产品供选择, 下面是生产商提供的产品性能指标:

印刷机型号	使用成本 (\$/小时)	印刷速度 (张/分钟)
1	15	30
2	20	36
3	24	50
4	27	66

KeenCo 公司的印刷业务按泊松分布到达, 平均每天 24 小时有 4 宗业务. 印刷量是随机的, 但平均每宗业务有 10 000 张. 与顾客签订的合同规定, 不能按期交活的罚款额为每宗业务每天 \$80. 那么 KeenCo 应该购买哪种印刷机呢?

令下标 i 代表印刷机型号 $i(i=1,2,3,4)$, 则相应于印刷机 i 每天的总期望费用为

$$\begin{aligned}ETC_i &= EOC_i + EWC_i \\ &= C_{1i} \times 24 + C_{2i} L_{si} \\ &= 24C_{1i} + 80L_{si}, i=1,2,3,4\end{aligned}$$

C_{1i} 的值由问题的数据给出. 我们知道, 在实际中的每台印刷机都可以看成是一个 $(M/M/1):(GD/\infty/\infty)$ 模型. 到达率为 $\lambda=4$ 宗业务/天. 求出相应于 i 型印刷机的服务率 μ_i 为

印刷机型号 i	服务率 μ_i(宗业务/天)
1	4.32
2	5.18
3	7.20
4	9.50

我们给出型号 1 的服务率计算过程.

$$\text{每宗业务的平均时间} = \frac{10\ 000}{30} \times \frac{1}{60} = 5.56 \text{ 小时}$$

因此有

$$\mu_1 = \frac{24}{5.56} = 4.32 \text{ 宗业务/天}$$

下表给出用 TORA 或 excelPoissonQ.xls 计算的 L_{si} 的结果:

印刷机型号 i	λ_i(宗业务/天)	μ_i(宗业务/天)	L_{si}(宗业务)
1	4	4.32	12.50
2	4	5.18	3.39
3	4	7.20	1.25
4	4	9.50	0.73

这 4 个型号印刷机的费用计算如下:

印刷机型号 i	EOC_i(\$)	EWC_i(\$)	ETC_i(\$)
1	360.00	1 000.00	1 360.00
2	480.00	271.20	751.20
3	**576.00**	**100.00**	**676.00**
4	648.00	58.40	706.40

其中型号 3 的费用最低.

习题 12.9A

1. 在例 12.9-1 中, 求下列各项:

(a) 计算出例中 μ_2, μ_3, μ_4 的值.

(b) 假定每天每宗业务 \$80 的延期赔偿款只适用于当天末还没有“开印”的业务, 则哪种型号的印刷机能给出每天最低总费用?

***2.** Metalco 公司正在为有 10 台机床的车间聘用一名修理工. 有 2 个候选人供考虑: 第 1 个候选修理工每小时能修好 5 台机床, 工资是每小时 \$15. 而第 2 个候选工人技术更好, 每小时工资 \$20, 能在一小时内修理 8 台机床. Metalco 估计, 每台出故障的机床因为耽误了生产, 每小时要损失费用 \$50. 假定机床的故障率服从泊松分布, 平均每小时 3 台, 并且修理时间是指数分布的, 那么该雇用哪个修理工呢?

3. B&K 食品公司开了一家新店, 号称用了“最先进的”收款台读码仪. Bih 先生是 B&K 的老板之一, 规定选择两种型号的读码仪之一. 读码仪 A 一分钟能处理 10 件商品, 而质量更好的读码仪 B 一分钟能处理 15 件商品. A 型和 B 型读码仪每天 (按 10 小时) 运行费和维护费用分别是 \$25 和 \$35. 买完东西的顾客按照泊松分布来到收款台, 到达率为每小时 10 位顾客. 每位顾客的购物车装有 25 件到 35 件商品, 呈均匀分布. Bih 估计, 每位等待顾客每分钟的平均费用约为 20 美分. 那么 B&K 店应该购买哪种读码仪呢?(提示: 每位顾客的服务时间不是指数分布, 而是均匀分布的.)

4. H&I 公司根据客户的不同要求生产一种可设定不同生产率 (每小时生产件数) 的专用机床. 一家加工厂的老板正在考虑购买一台这样的机床, 他看中的是最经济的加工速度. 根据过去的经验, 这位老板估计来自顾客的订单服从泊松分布, 每小时平均 3 份订单. 每份订单平均要货 500 件. 该老板和他的顾客之间有协议, 如果完不成, 每份订单每小时罚款 \$100.

(a) 假定每份订单的实际生产时间服从指数分布, 请建立一个作为生产率 μ 的函数的一般费用模型.

***(b)** 根据 (a) 的费用模型, 求最优生产率的表达式.

***(c)** 利用本问题给出的数据, 求该老板应该向 H&I 订货机床的最优生产率.

5. 某机械加工厂的工作按照泊松分布到达, 每周平均 80 件. 有一台自动机床成为该工厂的生产瓶颈. 根据估算, 该机床的生产率每增加一个单位, 每周将会节省费用 \$250. 不能按时完成的工作通常会导致业务的流失, 每周每件工作估计损失 \$500. 求该台自动机床的最优生产率.

6. 某比萨连锁企业销售两种规模的加盟餐馆模式. 模式 A 的规模能接待 20 拨顾客, 而模式 B 能为 30 拨顾客提供座位. 模式 A 的月经营成本为 \$12 000, 模式 B 为 \$16 000. 一个投资人想要开办一家自助式比萨饼店, 他估计顾客拨数 (每拨占一桌) 服从每小时 25 拨的泊松分布. 如果所有的桌子都有人了, 顾客就去别的店. 模式 A 每小时能接待 26 拨, 模式 B 每小时接待 29 拨. 因为每拨顾客的人数不同, 定餐的方式也不同, 所以服务时间是指数分布的. 该投资者预计, 流失业务的平均费用为每拨顾客每小时 \$15, 不能为等待顾客按时提供服务所造成的平均赔偿费用为每小时每拨顾客 \$10.

(a) 建立一个合理的费用模型.

(b) 假定餐馆每天营业 10 个小时, 该向投资人建议哪一种模式?

7. 在第 6 题中假设, 该投资人基于每张额外餐桌的具体边际成本选择合适的餐馆容量. 请给出相应的一般费用模型, 并定义所有的模型项.

8. 二手货物流通公司以委托销售方式销售流行的二手货物. 该公司的经营方式可以看成是一类库存问题, 按照每天 λ 和 μ 件速率的泊松分布补充和消耗库存. 缺货每单位时间, 公司就会由于丧失机会而损失 $\$C_1$; 而每单位时间存货, 需要库存费 $\$C_2$.

(a) 建立一个每单位时间期望总费用的公式.

(b) 求出 $\rho = \frac{\lambda}{\mu}$ 的最优值. 为了让求出的解与 $(M/M/1):(GD/\infty/\infty)$ 模型的假设相一致, 对于 C_1 和 C_2 的相对值应该满足什么样的条件?

例 12.9-2

在一家有多个员工的工具库, 交换工具的请求按照泊松分布发生, 每小时平均有 17.5 个请求. 每个员工每小时平均能办理 10 个请求. 工具库雇用一名新员工的工资是每小时 \$12, 每小时每台等待机床的生产损失成本约为 \$50. 求该工具库最优的员工数.

这个情形对应于 $(M/M/c)$ 模型, 其中需要求出 c 的最优值. 因此, 在本节开始给出的一般费用模型中, 我们让 $x = c$, 得出下面的费用模型:

$$\begin{aligned} ETC(c) &= C_1 c + C_2 L_s(c) \\ &= 12c + 50L_s(c) \end{aligned}$$

注意 $L_s(c)$ 是该工具库 (并列) 员工数的函数.

我们用每小时 $\lambda = 17.5$ 个请求和每小时 $\mu = 10$ 个请求的 $(M/M/c):(GD/\infty/\infty)$ 模型. 在这种情况下, 仅当 $c > \frac{\lambda}{\mu}$(即对本例中的 $c \geqslant 2$) 时, 该模型达到平稳状态. 下表给出了求最优 c 值的计算结果, $L_s(c)$ 的值 (用 excelPoissonQ.xls 或 TORA 算出) 表明, 最优员工数应该为 4.

c	$L_s(c)$(请求)	$ETC(c)$(\$)
2	7.467	397.35
3	2.217	146.85
4	**1.842**	**140.10**
5	1.769	148.45
6	1.754	159.70

习题 12.9B

1. 假设 $C_1 = \$20, C_2 = \$45$, 求解例 12.9-2.

***2.** Tasco 石油公司拥有一套连续运行的输油管线增压器装置. 每台增压器的故障间隔时间服从指数分布, 平均是 20 小时. 维修时间也服从指数分布, 平均维修时间为 3 小时. 在一个特殊的增压站, 由 2 名维修人员看管 10 台增压器. 每个维修人员的小时工资是 \$18. 输油管线的故障损失费用为, 每台增压器每小时损失 \$30. Tasco 公司正在研究是否要雇用另一名维修人员.

(a) 雇用第 3 个维修人员是否会节约总费用?

(b) 当有 2 个或 3 个值班的维修人员时, 每次故障的损失有多少?

3. 某公司租用一条广域通信服务 (WATS) 电话专线, 月租费为 $2 000. 公司办公室每月工作 200 个小时. 在所有其他时间里该 WATS 服务用于其他业务, 公司不能使用. 上班期的 WATS 线路可挂 100 个销售人员的分机, 每个人可能在任何时候使用这条线路, 但每天 8 小时内平均 2 次, 呼叫之间的间隔时间服从指数分布. 如果占线的话, 销售人员会一直等待, 由于不方便造成的等待费用约为每分钟 1 美分. 我们假定在销售人员等待用电话的时候, 没有其他的呼叫需求. 正常的电话费 (不用 WATS 专线) 约为每分钟 50 美分, 每次通话时间服从指数分布, 平均是 6 分钟. 该公司正在考虑 (按同样的价格) 租用第二条 WATS 线路, 以改善服务.

(a) 这一条 WATS 专线比没有 WATS 系统会为公司节约费用吗? 比起不用 WATS 系统能让公司节约或者损失多少钱?

(b) 公司应该租用第二条 WATS 专线吗? 比起一条 WATS 专线, 公司会节约或者损失多少钱?

***4.** 某个机械加工车间有 20 台机床和 3 个修理工人. 一台工作的机床会按照泊松分布随机地发生故障. 每台机床的修理时间服从指数分布, 平均为 6 分钟. 对这一情况的排队分析表明, 每天 8 小时整个车间平均报修为 57.8 次. 假设每台机床的生产率为每小时 25 个工件, 每加工一个工件的收益是 $2. 进一步假设, 每个修理工的每小时工资额为 $20. 请对雇用修理工的费用与机床故障导致的收益损失费用进行比较.

5. $ETC(c)$ 在 $c = c^*$ 达到最小值的必要条件为

$$ETC(c^*-1) \geqslant ETC(c^*) \quad \text{和} \quad ETC(c^*+1) \geqslant ETC(c^*)$$

说明这些条件可简化为

$$L_s(c^*) - L_s(c^*+1) \leqslant \frac{C_1}{C_2} \leqslant L_s(c^*-1) - L_s(c^*)$$

对例 12.9-2 利用该结果, 并证明 $c^* = 4$.

12.9.2　渴望水平模型

费用模型的适用性依赖于我们能否对费用参数做出很好的估计. 一般情况下, 这些参数是很难估计的, 特别是与顾客等待时间有关的参数. 我们希望通过直接用排队系统性能度量指标的渴望水平模型来减轻这个困难. 这一模型的思路是, 通过对冲突的性能度量指标设定一些合理的上限, 来求出服务水平 (μ 或 c) 可以接受的范围. 这些上限就是决策者想要达到的渴望水平.

下面来说明如何把这种方法用于多服务台模型, 其中我们要确定某种“可以接受的”服务台数 c^*. 考虑下面两个 (冲突的) 度量指标:

(1) 系统中的平均时间 W_s;

(2) 服务台的空闲百分率 X.

空闲百分率的计算如下:

$$X = \frac{c-\bar{c}}{c} \times 100 = \frac{c-(L_s-L_q)}{c} \times 100 = \left(1-\frac{\lambda_{\text{eff}}}{c\mu}\right) \times 100$$

(证明见习题 12.6E 的第 12 题.)

这个问题简化为求服务台数 c^*, 使得

$$W_s \leqslant \alpha, \quad X \leqslant \beta$$

其中 α 和 β 为决策者指定的渴望水平. 例如, 我们可以规定 $\alpha = 3$ 分钟, $\beta = 10\%$.

为了求解这个问题, 可以画出 W_s 和 X 作为 c 的函数的图像, 如图 12.10 所示. 在图像中找到 α 和 β, 我们马上就能确定 c^* 的可接受范围. 假如这两个条件不能同时满足, 则在得出可行性范围之前, 必须对其中一个范围或者两个范围都进行放松.

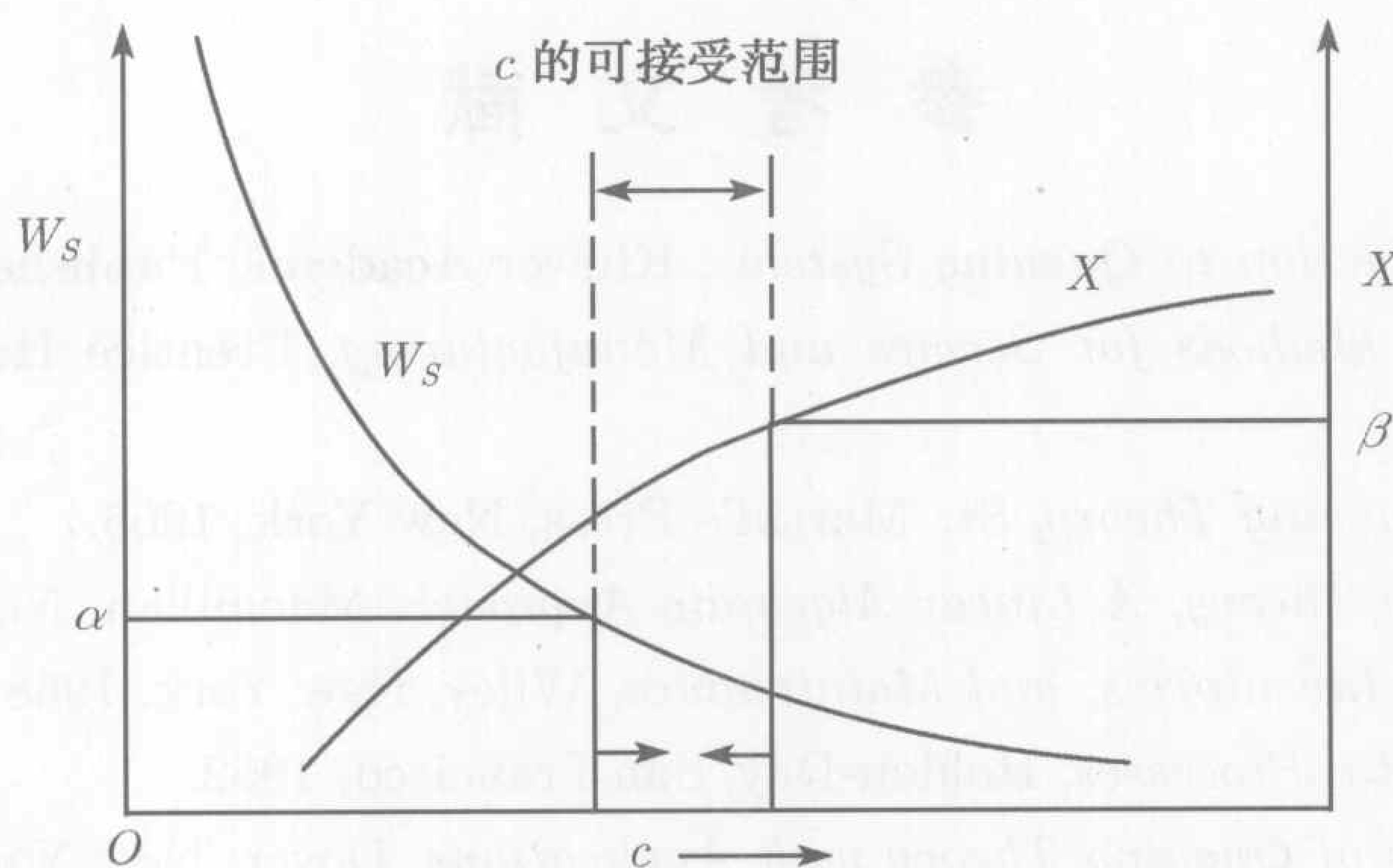


图 12.10 用渴望水平模型进行排队决策

例 12.9-3

在例 12.9-2 中, 假定我们需要确定员工人数, 使得得到一个工具的期望等待时间不超过 5 分钟. 同时, 还要求空闲百分率低于 20%.

在计算之前我们就知道, 在得到工具前等待时间为 5 分钟的渴望上限 (即 $W_s \leqslant$ 5 分钟) 显然是不合理的. 因为根据问题的数据, 平均服务时间本身就有 6 分钟. 下表给出了作为 c 的函数的 W_s 和 X 值.

c	2	3	4	5	6	7	8
W_s(min)	25.4	7.6	6.3	6.1	6.0	6.0	6.0
X(%)	12.5	41.7	56.3	65.0	70.8	75.0	78.0

根据这些结果, 我们应该减少服务时间, 或者注意到该问题的背景中, 工具的请求率高得不合理 ($\lambda =$ 每小时 17.5 个请求). 这很可能就是我要解决的问题. 例如, 我们可能想要研究一下工具替换率这么高的背后原因, 是不是工具的设计本身存在什么缺陷, 或者是因为机床操作工人的人为故意干扰因素?

习题 12.9C

*1. 一加工厂有 10 台同样的机器. 每台机器平均每 7 个小时发生一次故障. 修理一台故障机器平均需要半个小时. 故障和修理过程都服从泊松分布, 求:

(a) 要使得平均故障机器数不超过 1, 需要多少修理工?

(b) 需要多少修理工才能使得开始修理前的等待时间不超过 10 分钟.

2. 在 12.9.1 节的费用模型中, 一般很难估计出费用参数 C_2(等待费用). 因此, 可能需要用渴望水平的方法计算出费用 C_2. 用渴望水平模型求出 c^*, 我们就能用下面的不等式估计出可能的 C_2:

$$L_s(c^*) - L_s(c^*+1) \leqslant \frac{C_1}{C_2} \leqslant L_s(c^*-1) - L_s(c^*)$$

(公式推导见习题 12.9B 的第 5 题). 假定 $c^* = 3, C_1 = \$12.00$, 将这一方法用于例 12.9-2 的问题.

参 考 文 献

Bose, S., *An Introduction to Queuing Systems*, Kluwer Academic Publishers, Boston, 2001.

Hall, R., *Queuing Methods for Service and Manufacturing*, Prentice Hall, Upper Saddle River, NJ, 1991.

Lee, A., *Applied Queuing Theory*, St. Martin's Press, New York, 1966.

Lipsky, L., *Queuing Theory, A Linear Algebraic Approach*, Macmillan, New York, 1992.

Morse, P., *Queues, Inventories, and Maintenance*, Wiley, New York, 1958.

Parzen, E., *Stochastic Processes*, Holden-Day, San Francisco, 1962.

Saaty, T., *Elements of Queuing Theory with Applications*, Dover, New York, 1983.

Tanner, M., *Practical Queuing Analysis*, McGraw-Hill, New York, 1995.

Tijms, H.C., *Stochastic Models——An Algorithmic Approach*, Wiley, New York, 1994.

附录 A　AMPL 建模语言

本附录[①] 提供了求解与开发复杂数学规划模型所需的 AMPL 主要语法规则. 对于其他的细节, 可以参看附录后面所列的有关基本语言的参考文献 (Fourer 等人, 2003). 还可以参看网站 www.ampl.com 提供的附加资源、最新消息和更新资料.

A.1　初识 AMPL 模型

AMPL 采用普通书写格式为数学规划 (线性、整数和非线性规划) 建模提供了一个工作环境. 图 A.1 给出 Reddy Mikks 模型 (浅显易懂) 的线性规划程序 (文件 RM1.txt). 所有专用的关键词, 除了特殊的运算符 (+ − ∗ , ; : > < =) 以外, 全部用黑体. 其他的所有名称均由用户生成.

```
ampl: model RM1.txt;

    var x1 >=0;
    var x2 >=0;
    maximize z: 5*x1+4*x2;
    subject to
        c1: 6*x1+4*x2<=24;
        c2: x1+2*x2<=6;
        c3: -x1+x2<=1;
        c4: x2<=2;
    solve;
    display z,x1,x2;
```

图 A.1　初识 AMPL 模型 (文件 RM1.txt)

AMPL 在 DOS 环境下使用命令行和相关操作. 可以在 www.OptiRisk-Systems.com 网站上找到最近 Windows 交互界面的 beta 版测试软件.

可以点击 AMPL 文件夹中的**ampl.exe**文件来执行某一模型, 并根据**ampl**的提示, 输入后续命令, 并回车.

① 文件夹 AppenAFiles 在下载资料中, 它包括了本附录的全部文件.

其输出结果将在屏幕上[①] 有如下的显示:

```
MINOS 5.5: optimal solution found.
2 iterations
z = 21
x1 = 3
x2 = 1.5
```

这里给出了基本的普通书写格式, 但并不推荐此方法用于求解实际问题, 因为它是基于问题的特殊形式写成的. 本附录的其余内容将详细说明如何在实际应用中使用 AMPL.

A.2　AMPL 模型的组成

图 A.2 叙述 AMPL 模型的一般结构. 模型由两个基本段组成：顶段 (1 到 4) 是模型的代数表示, 底段 (5 到 7) 提供了驱动代数模型的相应数据. 因此, 在线性规划中, AMPL 代数表示形式完全相当于下面的数学模型:

代数表示	(1) 集的定义 (2) 参数的定义 (3) 变量的定义 (4) 模型表示 (目标和约束)
模型补充	(5) 输入数据 (6) 模型结果 (7) 结果输出

图 A.2　AMPL 模型的基本结构

$$\max \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \quad i = 1, 2, \cdots, m$$

这种安排的优点是, 相同的代数模型可以用于求解不同维数的线性规划问题, 只需简单地改变输入数据 m, n, c_j, a_{ij}, b_i 的值, 就可以求解任何维数的问题.

使用 AMPL 模型, 要注意以下一些语法规则:

① AMPL 的每种版本都有默认的规划求解, 它执行求 AMPL 模型最优解所需的计算. 在学生版中, MINOS 是默认的规划求解, 它可以处理线性和非线性问题. 下载资料中包含了其他的规划求解: CPLEX、KNITRO、LPSOLVE 和 LOQO. CPLEX 处理线性、整数和二次问题, LPSOLVE 处理线性和整数问题, KNITRO 和 LOQO 处理线性和非线性问题.

(1) AMPL 文件必须是纯文本的 (Windows 的记事本编辑器就能产生纯文本).

(2) 注释文字可以出现在模型的任何位置, 并用字符 # 引导.

(3) 每一条 AMPL 语句, 除注释语句外, 一定要以分号 (;) 结束.

(4) 一条 AMPL 语句可以超过一行, 断点最好选择在适当的位置, 如空格、冒号、逗点、圆括号、花括号、方括号, 或者是数学表达式. 这一规则对于字符串 (由单引号 '' 或双引号 "" 括起来的内容) 不适用, 此时, 需要在断点处加一个反斜杠 (\).

(5) 所有的关键词 (极少数除外) 均用小写字母.

(6) 用户使用的名称是区分大小写的. 如果必要, 其名称必须由字母、数字和下划线组成. 其他特殊的字符不能用于命名.

我们将用 2.1 节中的 Reddy Mikks 问题来描述 AMPL 的工作原理. 图 A.3 给出了相应的模型 (文件 RM2.txt). 为方便起见, 关键词 (或保留字) 都用黑体来强调.

代数模型以**集** (set) 开始, 它定义了一般线性规划模型的下标集. 集由用户命名, 如在 Reddy Mikks 模型中, 集的名称为 resource 和 paint. 每一个集必须由关键词 set 引导, 上面的两个集对应于一般线性规划模型中的下标集 $\{i\}$ 和 $\{j\}$. 该模型的输入数据段给出了集 resource 和 paint 的特定元素.

参数 (parameter) 由用户命名, 并由关键词 param 引导, 它们把目标函数和约束的系数定义为变量组和约束组的函数. 参数 unitprofit{paint}, aij{resource, paint}, rhs{resource} 分别对应于一般线性规划模型中的数学符号 c_j, a_{ij}, b_i. 下标 i 和 j 分别由 AMPL 中的集 resource 和 paint 来表示. 输入数据提供了参数的特定值.

模型的**变量** (variable) 是 x_j, 在这个模型中叫 product, 它由关键词var 引导. 同样, product 是集 paint 的函数. 可以在这同一条语句中增加非负限制 (>=0). 否则, 默认值表示变量无符号限制.

在模型的集、参数和变量的定义完成之后, 下一步是利用这些符号表示最优化问题. 目标函数语句的特定含义就是最优化, 所使用的关键词是 maximize 或者 minimize. 目标值 z 是由 profit 加冒号 (:) 表示, 它的 AMPL 语句

```
sum{j in paint}unitprofit[j]*product[j];
```

是数学表达式 $\sum_j c_j x_j$ 的直接转述. 下标 j 由用户指定. 注意, 在花括号 {j in paint} 中使用的 j, 表示 j 是集 paint 中的成员; 在方括号 [j] 中使用的 j, 表示 j 是下标.

一个模型可以包含一条或多条约束语句, 并且每条这样的语句都可以由关键词 subject to 或其简写形式 s.t. 引导. 实际上, s.t. 和subject to写不写均可, 因为 AMPL 规定, 任何没有以关键词作为开始的语句都是约束. 在 Reddy Mikks

```
#*******************ALGEBRAIC MODEL******************
#-------------------------------------------------sets
set paint;
set resource;
#-------------------------------------------parameters
param  unitprofit{paint};
param  rhs {resource};
param  aij {resource,paint};
#--------------------------------------------variables
var product{paint} >= 0;
#------------------------------------------------model
maximize profit: sum{j in paint} unitprofit[j]*product[j];
subject to limit{i in resource}:
      sum{j in paint} aij[i,j]*product[j] <= rhs[i];
#***********************DATA************************
data;
set paint := exterior interior;
set resource := m1 m2 demand market;
param unitprofit :=
                exterior 5
                interior 4;
param rhs:=
                m1  24
                m2  6
                demand  1
                market  2;

param aij:  exterior  interior :=
          m1        6    4
          m2        1    2
          demand   -1    1
          market    0    1;

#*********************SOLUTION**********************
solve;
#---------------------------------------output results
display  profit, product, limit.dual, product.rc;
```

图 A.3 Reddy Mikks 问题的 AMPL 模型 (文件 RM2.txt)

模型中, 只有一组名叫 limit 的约束, 它和集 resource 有关:

```
limit{i in resource}:
        sum{j in paint} aij[i,j]*product[j] <= rhs[i];
```

这条语句是第 i 个约束 $\sum_j a_{ij}x_j \leqslant b_i$ 的直接转述.

声明变量是非负的这一思想可以推广到允许建立变量的上下界. 因此, 可以去掉那些需要声明界的约束. 首先, 声明下界和上界, 不妨起名为 lowerbound 和 upperbound, 如

```
param lowerbound{paint};
param upperbound{paint};
```

其次, 变量被定义为

```
var product{j in  paint}>=lowerbound[j],<=upperbound[j];
```

注意, 语法规则不允许比较 "向量". 因此, 如果使用

```
var product{paint}>=lowerbound{paint},<=upperbound{paint};
```

将会产生语法错误.

也可以使用相同的语法设置参数条件, 例如, 语句

```
param upperbound{j in paint}>=lowerbound[j];
```

将保证 upperbound 不小于 lowerbound, 否则 AMPL 将报错. 对参数使用界的主要目的是, 防止无意中输入相互矛盾的数据. 另外, 当参数只需要假定非负值时, 可以使用这种检查方法.

图 A.3 中的代数模型具有普遍性, 它可以应用于任意多个变量和任意多个约束. 可以通过给出该问题的特定数据, "量体裁衣" 式地应用于 Reddy Mikks 情形. 在语句 data; 之后, 我们首先定义集的成员, 然后再用这些定义为不同参数赋值.

集 paint 包含两个变量的名称, 这里我们称之为 exterior 和 interior. 集 resource 的成员由名叫 m1、m2、demand 和 market 的变量组成. 因此, data 段给出相应的语句是

```
set paint  := exterior interior;
set resource := m1 m2 demand market;
```

每个集的成员出现在保留运算符 := 的右端, 并由空格 (或逗号) 隔开. 当数据段外部 调用它们时, 字符串指标必须包含在双引号中, 即 paint["exterior"]、paint["interior"]、limit["m1"]、limit["m2"]、limit["demand"] 和 limit["market"]). 否则, 字符串指标会被错误地解释成 (数值型) 参数.

如果在代数模型的开始处 (而不是在数据段中), 我们已定义了如下的集:

```
set resource = {"m1", "m2", "demand", "market"};
set paint = {"exterior", "interior"};
```

(注意, 双引号 ""、分隔逗号和花括号都是不可少的.) 那么这种方法对于一般线性规划问题并不可取, 因为它针对的是本问题, 这限制了本模型用于不同的输入数据情形中. 使用这种方法, AMPL 将不再允许在 data 段中修改集的成员.

在集 resource 和 paint 的成员中, 使用字母数字的命名方式对于大问题可能过于繁琐. 因此, AMPL 允许使用纯数字集, 即可以使用

```
set paint:= 1 2;
set resource:= 1..4;
```

范围 1..4 代替了明确的 1 2 3 4 的表达方式, 这对于成员数目很大的集非常有用. 例如, 1..1000 表示的是具有 1 000 个成员的集, 成员以 1 为开始, 以 1 000 为结束, 成员之间逐个加 1.

范围的表达可以采用更一般的形式, 首先定义 m 和 n 作为参数

```
param m;
param n;
```

在这种情况下, 集 1..m 和 1..n 可以直接用在整个模型中, 如图 A.4 所示 (文件 RM2a.txt), 完全不必使用集的名称 resource 和 paint.

```
param m;
param n;
param  unitprofit{1..n};
param  rhs{1..m};
param  aij{1..m,1..n};
#-----------------------------------------------variables
var product{1..n}>= 0;
#---------------------------------------------------model
maximize profit:sum{j in 1..n}unitprofit[j]*product[j];
subject to limit{i in 1..m}:
     sum{j in 1..n}aij[i,j]*product[j]<=rhs[i];
data;
param m:=4;
param n:=2;
param unitprofit := 1 5  2 4;
param rhs:= 1 24  2 6  3 1  4 2;
param aij: 1    2:=
        1  6    4
        2  1    2
        3  -1   1
        4  0    1;
solve;
display profit, product, limit.dual, product.rc;
```

图 A.4 Reddy Mikks 问题的 AMPL 模型 (文件 RM2a.txt)

实际上, 语法 1..m(或 1..n) 有更一般的形式

```
start..end by step
```

其中 start、end 和 step 是由 AMPL 定义的参数, 它的值在 data 后面指定. 如果 start<end 并且 step>0, 则集的成员由 start 开始, 其增加值为 step, 直到增加到小于或等于 end 的最大值停止. 如果 start>end 并且 step<0, 则是相反的情况. 例如, 3..10 by 2 产生的成员是 3, 5, 7, 10..3 by -2 产生的成员是 10, 8, 6, 4. step 的默认值为 1. 这意味着start..end by 1与 start..end 是相同的.

实际上, 在任何执行时, 参数 start、end 和 step 可以是任意的合法的 AMPL 数学表达式. 例如, 已知参数 m 和 n, 集 j in 2*n..m+n^2 by n/2 是完全合法的. 然而, 注意到分数的 step 可以直接用于产生集的成员. 例如, 对于 m=5, n=13, 则 m..n step m/2 的集成员是 5, 7.5, 10, 12.5.

Reddy Mikks 模型中包含一维和二维参数. 参数 unitprofit 和 rhs 属于前一类, 而参数 aij 属于后一类. 在第一类参数中, 通过在每个集成员后面指定一个数值而指定其数据, 其表达形式如下:

```
param unitprofit :=
      exterior  5
      interior  4;
param rhs:=
      m1        24
      m2        6
      demand    1
      market    2;
```

如果希望列表中的元素可以被"排成"一行, 那么元素之间只需用空格隔开. 值得提倡的是目前这种格式, 因为它有更好的可读性.

二维参数 aij 的数据输入类似于一维情况. 不同的是, 在 aij: 后必须指定列的次序, 以消除可能产生的混淆, 如以下的表达方式所示:

```
param aij:             exterior  interior :=
       m1                 6         4
       m2                 1         2
       demand            -1         1
       market             0         1;
```

列表同样可以是自由格式, 只要保证相应的次序是符合逻辑的, 并且其元素由空格隔开.

AMPL 允许为参数的所有元素指定默认值. 例如, 假定对于参数 c, 有 $c_1 = 11$, $c_8 = 22$, 而且 $c_i = 0$, $i = 2, 3, \cdots, 7$. 可以用下面的语句指定数据 c:

```
param c{1..8};
.
.
.
data;
param c:=1 11 2 0 3 0 4 0 5 0 6 0 7 0 8 22;
```

得到同样结果的更紧凑的方法是使用下面的语句:

```
param c{1..8} default 0;
.
.
.
data;
param c:=1 11 8 22;
```

最初, c[1] 到 c[8] 均取默认值 0, 然后, c[1] 和 c[8] 再在 data 段中被改为 11 和 22. 通常, default(默认值) 的后面可以跟任何数学表达式. 这个表达式在程序开始时仅被赋值一次.

AMPL 模型的最终段用于得到解和表达输出. 命令 solve 是所有程序求解模型所必需的. 一旦完成后, 可以请求指定输出结果. 跟随命令 display 后面的输出列表只是一种显示结果的方法. 在 Reddy Mikks 模型中, 语句

```
display profit, product, limit.dual, product.rc;
```

显示目标函数 profit、变量 product、约束的对偶 limit.dual 和变量的简约费用 product.rc 的最优值. 关键词 dual 和 rc 分别是约束 limit 和变量 product 的后缀, 用点 (.) 隔开. 它们不能作为单独的关键词使用. 输出的默认值是屏幕. 也可以直接输出到外部文件中, 其方法是在分号 (;) 前直接插入 >文件名. 附录 A.5 将详细介绍如何直接将输出送到文件和电子表格中.

DOS 下的执行命令是

```
ampl: model RM2.txt;
```

相应的输出显示在屏幕上, 如图 A.5 所示.

在图 A.5 中, 输出的布局安排有点 "杂乱", 因为它将约束指标和变量指标混在一起. 可以重新组织输出的形式, 通过使用下面两条 display 语句, 将它的元素放在相同维数的数组中, 如:

```
display profit, product, product.rc;
display limit.dual;
```

在一个典型的 AMPL 模型中, 正如图 A.3 所示的模型那样, 模型的逻辑段更适合保持静态. 数据和输出段需要作相应的改变, 使之与特定的线性规划问题相匹

```
MINOS 5.5: optimal solution found.
2 iterations, objective 21
profit = 21

:          product limit.dual   product.rc    :=
demand       .       0           .
exterior     3       .          6.66134e-16
interior    1.5      .           0
m1           .      0.75         .
m2           .      0.5          .
market       .       0           .
;
```

图 A.5 在 Reddy Mikks 模型中，使用display profit, product, limit.dual. product.rc的 AMPL 输出

配. 为此, AMPL 模型可以由两个独立的文件表示：文件 RM2b.txt 提供模型的逻辑部分, 而文件 RM2b.dat 提供输入数据, 并说明输出结果[①]. 在这种情况下, DOS 行命令可按下列次序输入：

```
ampl: model RM2b.txt;
ampl: data RM2b.dat;
```

我们将在附录 A.7 中看到, 像 solve 和 display 这样的命令可以采用交互式的方式来输出, 而不是模型中的硬编码.

Reddy Mikks 模型展示了 AMPL 众多能力中很少的一点. 在后面, 我们将展示如何从外部文件和电子表格中读取输入数据. 我们还将展示, 如何向这些媒体传送专门设计的 (有格式的) 输出. 还有, AMPL 交互式命令是重要的调试和执行工具, 这一点将在附录 A.7 作进一步的解释.

习题 A.2A

1. 修正图 A.3 所示的 Reddy Mikks AMPL 模型 (文件 RM2.txt), 考虑具有防水功能的涂料在内共有 3 种涂料. 每吨防水涂料需要原料 m1 和 m2 分别为 0.5 吨和 0.75 吨. 新涂料的日需求量在 0.5 吨至 1.5 吨之间. 每吨的收入是 3 500 美元. 这种产品没有其他的限制.
2. 在图 A.3 所示的 Reddy Mikks AMPL 模型 (文件 RM2.txt) 中, 用下列集的定义重写 AMPL 程序：

   (a) paint 和 {1..m}　　(b) {1..n} 和 resource

   (c) {1..m} 和 {1..n}
3. 修正图 A.3 所示的 Reddy Mikks 模型 (文件 RM2.txt) 中变量的定义, 使之包含外墙涂料的最小需求量 1 吨, 外、内墙涂料的最大需求量分别是 2 吨和 2.5 吨.

① 实际上, 输出命令可单独执行, 而不是包含在.dat 文件中, 这一点将在附录 A.7 中作进一步的解释.

4. 在图 A.3 所示的 Reddy Mikks AMPL 模型中, 命令

```
display profit;
```

提供了目标函数值. 可以用相同的命令去显示每个变量对总利润的贡献, 其命令格式如下:

```
display profit, {j in paint} unitprofit[j]*product[j];
```

另一种完成相同结果的方便方法是使用**定义变量** (defined variable) 语句, 其格式如下:

```
var  extProfit=unitprofit["exterior"]*product["exterior"]
var  intProfit=unitprofit["interior"]*product["interior"]
```

在这种情况下, 目标函数和显示语句就可以写得简单一些, 其形式如下:

```
maximize profit: extProfit + intProfit;
display profit, extProfit, intProfit;
```

事实上, 定义变量可以有其指标形式, 如:

```
var varProfit{j in  paint} = unitprofit[j]*product[j];
```

这样得到的目标函数和显示语句可以写成如下形式:

```
maximize profit: sum {j in  paint} varProfit[j];
display profit, varProfit;
```

在 Reddy Mikks 模型中, 使用定义变量允许显示每个变量的利润贡献和原料 m1 和 m2 的资源消耗.

5. 开发和求解例 2.2-2 营养配方问题的一个 AMPL 模型, 并求最优解. 确定和解释相应的对偶值和简约费用.

A.3　数学表达式和计算参数

我们已经看到, AMPL 允许设定参数的上下界. 实际上, AMPL 语言提供了更加灵活的方式, 如果希望的话, 可以把参数定义为复杂的数学表达式或者计算参数.

为说明计算参数的使用, 考虑银行提供 n 种贷款的情况, 其中第 i 种贷款的利息为 r_i, $0 < r_i < 1, i = 1, 2, \cdots, n$. 对于本金和利息不能全部收回的坏账, 设贷款 i 的比例为 v_i. 目标函数是确定银行分配贷款 i 的数量 x_i, 使得在约束条件限制下总收益最大.

说明计算参数的使用, 重点要集中在目标函数上. 目标函数的代数表达式可表示为

$$\max \quad z = \sum_{i=1}^{n} r_i(1 - v_i)x_i - \sum_{i=1}^{n} v_i x_i = \sum_{i=1}^{n} [r_i - v_i(r_i + 1)]\, x_i$$

直接将 z 转述成 AMPL 语句有如下形式:

```
param r{1..n}>0, <1;
param v{1..n}>0, <1;
var x{1..n}>=0;
maximize z: sum{i in 1..n}(r[i]-v[i]*(r[i]+1))*x[i];
(constraints)
```

处理银行情况的另一种方法是, 使用计算参数来表示目标函数的系数, 其格式如下:

```
param r{1..n}>0, <1;
param v{1..n}>0, <1;
param c{i in 1..n}=(r[i]-v[i]*(r[i]+1));
var x{1..n}>=0;
maximize z: sum{i in 1..n}c[i]*x[i];
(constraints)
```

AMPL 对参数 `c[i]` 作计算, 并将其值用在目标表达式 `z` 中. 新的公式增强了可读性. 但在某些情况中, 计算参数的使用可能是最基本的.

通常, 确定计算参数值的表达式可以任意复杂, 并且可以包含任何内置的算术函数, 这些内置函数是大家在任何程序语言中所熟知的 (例如, `sin`、`max`、`log`、`sqrt`、`exp`). 一项重要的要求是表达式的值是数值型的[①].

可以使用条件结构

```
参数 = if 条件 then 表达式 1 else 表达式 2;
```

对计算参数赋值. 条件作算术值比较, 其字符串也使用熟知的运算符 `=`、`<`、`>`、`<=`、`>=` 和 `<>`(甚至还有 `and/or`). 如要条件是模型变量的函数, 将会导致非线性. 像在其他程序语言一样, 其结构可以在没有 `else` *表达式* 2 的情况下使用. 在 `then` 和 `else` 的后面允许嵌套使用 `if` 语句.

`if-then-else` 结构赋给计算参数的值或者是*表达式* 1 或者是*表达式* 2. 这就是为什么 `if-then-else` 结构是*表达式*而不是*语句*的原因. (附录 A.7 将介绍 `if-then-else`*语句*, 以及循环语句 `for{}`、`repeat while{}` 和 `repeat until{}`. 这些语句主要用在自动解设定和格式化输出.)

这里用一个简单的例子来说明 `if` 表达式的使用. 例如在多周期生产情形中, 生产某种产品的件数要适应需求的变化. 估计每件产品的费用在第一个 m 周期是 p 美元, 下一个 m 周期增加 10%, 再下一个 m 周期增加 20%.

这个模型的约束是每个周期生产量的限制, 以及库存、产品和需求的平衡方程.

---

① AMPL 手册提供了一个 "例外". 当声明参数是*二元变量*时, 参数也可以看成是*逻辑型*的. 但这种差别是人为的, 因为将这一类参数按数值型来处理, 依然产生同样的结果.

为了说明 if 表达式的使用方法, 我们将重点集中在问题的目标函数上. 令

$$x_j = \text{周期 } j \text{ 生产产品的件数}, \quad j = 1, 2, \cdots, 3m$$

目标函数如下:

$$\begin{aligned}\min \quad z = & p(x_1 + x_2 + \cdots + x_m) + 1.1p(x_{m+1} + x_{m+2} + \cdots + x_{2m}) \\ & +1.2p(x_{2m+1} + x_{2m+2} + \cdots + x_{3m})\end{aligned}$$

可以在 AMPL 中建立函数模型如下:

```
param p;
var x{1..3*m}>=0;
minimize cost: p*(sum{j in 1..m}x[j]+1.1*sum{j in m+1..2*m}x[j]+
     1.2*sum{j in 2*m+1..3*m}x[j]);
(constraints)
```

增强可读性并采用一种较为简洁的形式, 方法是使用 if-then-else 来表示目标函数的参数 c[j]:

```
param m;
param n=3*m;
param p;
param c{j in 1..n}= if j<=m then p else
                        (if j>m and j<=2*m then 1.1*p  else 1.2*p);
var x{j in 1..n};
minimize z: sum{j in 1..n}c[j]*x[j];
(constraints)
```

注意条件的嵌套. 虽然不必写包含第 2 个 if 的圆括号 (), 但使用它可以增强语句的可读性. 我们注意到, then 和 else 后面总是跟着数值型计算表达式. 注意, c 也可以定义为

```
param c{j in 1..n}=p*( if j<=m then 1 else
                        (if j>m and j<=2*m  then 1.1 else 1.2));
```

if-then-else 的一种特别有用的使用方法是, 其参数或变量由递归形式定义. 这种参数的一个典型例子是在确定库存水平中, 其中周期 t 的库存水平为 I_t, $t = 1, 2, \cdots, n$, 且初始库存为 0. 周期 t 的生产量和需求量分别是 p_t 和 d_t. 因此, 库存水平是

$$I_0 = 0$$

$$I_t = I_{t-1} + p_t - d_t, \quad t = 1, 2, \cdots, n$$

库存水平 I_t 可以在 AMPL 中用递归的方式计算, 方法如下:

```
param p{1..n};
param d{1..n};
var I{t in 1..n}= if i=1 then 0 else I[t-1]+p[t]-d[t];
```

注意, 如果不使用 if-then-else 表达式, 计算 I_t 将会产生一些麻烦 (见习题 A.3A).

习题 A.3A

*1. 考虑下列约束:

$$\begin{aligned} x_i + x_{i+1} &\geqslant c_i, \quad i = 1, 2, \cdots, n-1 \\ x_1 + x_n &\geqslant c_n \end{aligned}$$

用 if-then-else 结构开发一组约束, 用来表示全部 n 个不等式.

2. 在多周期的产品库存问题中, 令 x_t, z_t, d_t 分别为周期 $t(t = 1, 2, \cdots, T)$ 的入库量、生产量和需求量. 与周期 t 相应的平衡方程为 $x_t + z_t - d_t - x_{t+1} = 0$. 在特定情况下, $x_1 = c(> 0)$, 且 $x_{T+1} = 0$. 运用 if-then-else 结构, 写出对应于平衡方程的 AMPL 约束, 同时包括 $x_1 = c$ 且 $x_{T+1} = 0$.

A.4 子集和指标集

子集 假设有下列约束:

$$x_1 + x_2 + x_5 + x_6 + x_7 \leqslant 15$$

模型中有 7 个变量, 但这个特殊约束不包含变量 x_3 和 x_4.

可以用不同方式使用子集来写出这个约束的模型 (所有新的关键词均用黑体).

```
#---------------------- method 1 --------------------------------
var x{1..7}>=0;
subject to lim: sum{j in 1..7: j<=2 or  j>=5}x[j]<=15;
#---------------------- method 2 --------------------------------
var x{1..7}>=0;
subject to lim: sum{j in 1..2 union 5..7}x[j]<=15;
#---------------------- method 3 --------------------------------
var x{1..7}>=0;
subject to lim: sum{j in 1..7 diff 3..4}x[j]]<=15;
#---------------------- method 4 --------------------------------
var x{1..7}>=0;
subject to lim: sum{j in 1..7 diff (1..4 inter  3..7)}x[j]]<=15;
#----------------------------------------------------------------
```

方法 1 中, 通过对 j 强加限制以从集 {j in1..7} 中删去元素 3 和 4. 冒号 (:) 将修正集与条件分开. 关键词 union、diff 和 inter 分别起着 $A \cup B$、$A-B$ 和 $A \cap B$ 的作用. 方法 4 中, 集的表示是令人费解的. 不过, 它提供了运算 inter 的一种使用方法.

指标集　AMPL 强有力的特色是允许指标集优先于普通集的元素. 假设用两种成分 A 和 B 来生产产品 1, 2, 3, 4, 5. 用成分 A 生产产品 1, 3, 5, 用成分 B 生产产品 1, 2, 4, 5. 每种产品需要一个单位的指定成分. 成分 A 和成分 B 的最大可用量分别是 200 和 300 单位. 问题是确定每种产品的装配单位数. 还需要其他相关的数据来完整地描述此问题, 但我们的重点仅集中在如何处理可利用成分的约束上.

令 x_i 是产品 $i(i=1,2,\cdots,5)$ 的生产量. 则成分 A 和成分 B 的约束的数学表达式可以表示成

$$\text{成分 A:}\ x_1 + x_3 + x_5 \leqslant 200$$

$$\text{成分 B:}\ x_1 + x_2 + x_4 + x_5 \leqslant 300$$

约束的 AMPL 表示可以使用指标集来实现, 其形式如下:

```
set comp;
set prod{comp};
param d{comp};
var x{1..5}>=0;
#-----objective function here
subject to
      C{i in comp}:sum{j in prod[i]}x[j]]<=d[i];
#-----other constraints here
data;
set comp:= A B;
set prod[A]:=1 3 5;
set prod[B]:=1 2 4 5;
param d:= A 200 B 300;
```

集 prod 的指标是集 comp 的元素 A 和 B, 因此定义两个指标集 prod[A] 和 prod[B]. 接下来, 该问题的数据段 (在 data; 后) 定义了 prod[A] 和 prod[B] 的元素. 结合这些数据, 关于成分的约束 (不管有多少个) 由单一语句定义:

```
C{i in comp}:sum{j in prod[i]}x[j]]<=d[i];
```

在例 6.5-4 和例 8.1-2 后面的 AMPL 程序中, 恰当地表现了指标集的应用.

习题 A.4A

1. 仅仅借助于 sum{} 函数, 用子集表示下面约束的左端项:

(a) $\sum_{j=1}^{m} x_j + \sum_{j=m+k}^{n} x_j + \sum_{j=n+p}^{q} x_j \geqslant c$

(b) $\sum_{i=m}^{n} x_i + \sum_{i=n+k}^{2n+k} x_i \leqslant c, \quad k \geqslant 1$

*2. 假设使用 5 种组件 (每件产品使用一个单位) 生产 10 种产品, 其相关数据如下表所示:

组件	使用该组件的产品	最小可利用量
1	1, 2, 5, 10	500
2	3, 6, 7, 8, 9	400
3	1, 2, 3, 5, 6, 7, 9	900
4	2, 4, 6, 8, 10	700
5	1, 3, 4, 5, 6, 7, 9, 10	100

每件产品的组装成本是所用组件的函数, 从组件 1 到组件 5 分别是 9 美元、4 美元、6 美元、5 美元和 8 美元. 任意一种产品的最大需求量是 300 件. 使用 AMPL指标集 确定最小化安装成本的最优产品组合. (提示: 令 x_{ij} 是使用组件 j 产品 i 的件数.)

3. 重做第 2 题, 假定组件的单位安装成本是所组装的产品的函数, 从产品 1 至产品 10 分别是 1 美元、3 美元、2 美元、6 美元、4 美元、9 美元、2 美元、5 美元、10 美元和 7 美元.

A.5 存取外部文件

到目前为止, 我们已使用 "硬编码" 数据驱动了 AMPL 模型. 实际上, 可以从外部文件、电子表格和/或数据库中读取 AMPL 的数据. 同样也可以将输出写到外部文件、电子表格和数据库中. 本节介绍如何从下列媒介读输入数据或写输出数据:

(1) 外部文件, 包括屏幕与键盘;　　(2) 电子表格.

更详细的资料可查阅参考文献 [Fourer 等人, 2003] 的第 10 章.

A.5.1 简单读文件

从无格式外部文件中读数据的语句是

read *项目列表* < *文件名*;

*项目列表*是由逗号分隔的无指标或有指标参数列表. 在有指标的情况下, 其语法规则是{*指标集*}*参数名* [*指标*]. 列表可以仅包含参数. 这意味着, 数据在调用 read 语句之前, 必须说明所有集的成员. (附录 A.5.3 和附录 A.5.4 将介绍, 集成员如何从格式文件和电子表格中读取.)

为解释 read 语句的用法, 考虑 Reddy Mikks 模型, 其中参数 unitprofit、rhs 和 aij 的所有数据从文件 RM3.dat 读出, 其模型在文件 RM3.txt 中. 相应的 read 语句是:

```
read {j in paint}unitprofit[j],
     {i in resource}rhs[i],
     {i in resource, j in paint}aij[i,j]<RM3.dat;
```

文件 RM3.dat 所列的数据与 read 列表中出现的项应有完全相同的顺序:

```
5    4
24   6    1    2
6    4
1    2
-1   1
0    1
```

数据排成多行是为了提高可读性, 从某种意义来说, 我们已有过将所有的元素排成一行 (元素之间用空格隔开) 的情况[①]. 注意, 这个文件恰好全是数值型数据. 为了方便起见, 非数值型的数据 (如参数的名称) 也可以出现在数据文件中, 只要它们被 symbolic 声明过 (细节请见参考文献 [Fourer 等人, 2003] 的 7.8 节和 9.5 节).

read 语句允许从键盘上读取数据. 在这种情况下, 文件名被减号所取代 —— 即用 <-. 在这种情况下, read 执行后将产生 DOS 提示符 ampl?, 并将不断地重复, 直到所有需要由 read 读取的数据全部读完为止.

习题 A.5A

1. 为 Reddy Mikks 模型 (文件 RM3.txt) 准备一个输入文件 RM3x.dat, 假定相应的 read 语句如下:

```
read {j in paint}
    {i in resource}
    (
    rhs[i],
    {j in paint}aij[i,j]
    )<RM3x.dat;
```

***2.** 对于 Reddy Mikks 模型, 解释为什么下面的 read 语句是繁琐的:

```
read {i in resource}
    (
    rhs[i],
    {j in paint}(unitprofit[j],aij[i,j])
    )<RM3xx.dat;
```

---

① .dat文件 (和本节后面介绍的.tab文件) 中的隐藏码可能引发 AMPL 错误, 如 “too few elements in line xx”(xx 行缺少元素, 其中 xx 是行的数值) 或 “unexpected end of file”(文件没有结束), 即使文件看起来似乎是完全合法的. 为了除去这些隐藏码, 立即单击文件中最近的数据元素的右端, 相继按回车键、退格键和回车键.

A.5.2 用 print 或 printf 将输出写到文件

在 AMPL 中, 写输出数据文件的一种简单的方法是用预定格式 print 或格式 printf. 作为解释, 看一下 Reddy Mikks 模型, 可以用下面的语句将输出数据传送到一个我们所命名的 file.out 文件中 (如果没有指定文件名, 输出的默认值是屏幕).

```
printf "Objective value is %6.2f\n",profit >file.out;
printf {j in paint}:
    "%8s%8.2f%8.3f\n",j,product[j],product[j].rc >file.out;
```

输出格式总是先于输出列表而且必须用双引号括起来. 去掉格式代码, 相同的语句可以出现在 print 中.

在第 1 个 printf 语句中, 格式包含选择项描述性文本 Objective value is, 以及如何输出列表的强制说明. 代码 %6.2f 说的是 profit 的值使用的格式具有 6 位长度, 其中小数点后有 2 位. 代码 \n 是移动文字到文件的下一行 (即相当于回车键). 这些格式代码与 C 程序是相同的.

在第 2 个 printf 语句中, 输出列表包含 j、product[j] 和 product[j].rc, 其中 j 是 AMPL 集 paint 中的成员 (exterior 和 interior) 之一. 代码 %8s 是为需要显示的名称 exterior 或 interior 预留的 8 个位置. 万一 j 是数值型 (如 {j in 1..2}), 则格式说明必须是整数型, 如 %3i.

本节的格式说明仅限于 %s、%i、%f 和 \n. AMPL 还提供了其他的说明 [见文献 (Fourer 等人, 2003) 的表 A-10].

习题 A.5B

1. 用 printf 语句描述 Reddy Mikks 模型 (文件 RM2.txt) 的最优解, 并使结果具有如下形式, 这里用后缀 .slack 和 .dual 得到松弛变量和对偶价格.

Objective value =

Product	Quantity	Profit($)
⋮	⋮	⋮
Constraint	**Slack amount**	**Dual price**
⋮	⋮	⋮

A.5.3 输入表文件

在附录 A.5.1 中, read 语句不允许读集成员. 这种情形可通过 table 语句来解决.

在 table 文件中, 数据以表格形式出现, 用所定义的集的成员适当地标记行和列. 读取 table 文件需要一个相伴的 read 语句. table 语句编排数据的格式, read

语句使这些数据在模型中可用.

table 语句和 read 语句的语法如下:

table *tableName* **IN**"*fileName*": *SetName* < −[*SetColHdng*], *parameters*∼*ParamColHdng*;

read table *tableName* ;

这种语法允许从*fileName* 的*tableName* 中读 AMPL 集的成员和参数成员.

默认的 *fileName* 纯文本表格是 *tableName* .tab. 可以使用明确指定的带有强制的.tab 扩展名的 *fileName*(在双引号中) 来覆盖它, 指定的 *fileName* 的位置在关键词 IN 后面. IN (大写字母) 意味着输入(对应的有 OUT, 将在后面介绍, 它用于把数据输出到表格文件). *SetColHdng* 可以是表中任意的一个标题名, 用 <- 交叉引用表中*SetName* 的元素. 类似地, 用 ∼ 交叉引用 AMPL 参数和任意的*ParamColHdng*. 如果 *SetColHdng* 与*SetName* 恰好相同, 则语法 *SetName* <-[*SetColHdng*] 可以被 [*SetName*] IN 代替. 至于参数, ∼*ParamColHdng* 可以从语句中删除.

为了解释表格的使用, 图 A.6 给出了名叫 RM4profit.tab、RM4rhs.tab 和 RM4aij.

File RM4profit.tab:

```
ampl.tab  1  1
COL1            COL2
exterior        5
interior        4
```

File RM4rhs.tab:

```
ampl.tab  1  1
resource        rhs
m1              24
m2              6
demand          1
market          2
```

File RM4aij.tab:

```
ampl.tab  2  1
resource   paint       aij
m1         exterior    6
m1         interior    4
m2         exterior    1
m2         interior    2
demand     exterior    -1
demand     interior    1
market     exterior    0
market     interior    1
```

图 A.6 表文件的内容, 用于输入 Reddy Mikks 模型的参数unitprofit、rhs和aij

tab 文件的内容, 这些文件用于输入 Reddy Mikks 模型的参数 unitprofit、rhs 和 aij. 每个文件的第一行必须是下面的格式:

ampl.tab *nbr_indexing_sets* *nbr_read_parameters*

第一个元素 ampl.tab 说明表格是一个 .tab 文件, 并且随后的两个元素提供将要从表中读出的参数指标集的数目. 在 RM4profit.tab 和 RM4rhs.tab 中, 只需一个集来定义参数 unitprofit 和 rhs, 为此, ampl.tab 1 1 用作这两个文件的标题行. 参数 aij 需要两个集来定义, 所以需要在标题行使用 ampl.tab 2 1.

标题行下面列出了集和参数的精确名称或替代名称. 在各自的文件中, 随后的行把输入参数的值列为其指标集明确的函数, 并用空格作为分隔符. 对于 unitprofit 和 rhs, 列表是很直观的. 对于双指标的参数 aij, 由两个明确的指标来说明每个参数列表, 即使是显得有些繁冗.

对于 Reddy Mikks 模型, 相应的表格定义如下:

```
table RM4profit IN: paint <- [COL1], unitprofit~COL2;
table RM4rhs IN: [resource] IN: rhs;
table RM4aij IN: [resource,paint], aij;
```

在 table 语句声明后, 可以用下面的语句读数据:

```
read table RMprofit;
read table RMrhs;
read table RMaij;
```

为了增加可读性, 建议把表格声明语句放在约束段后面. 于是, read 语句立刻放置在表格声明的下面 (见文件 RM4.txt).

上述表格语句解释了 4 条语法规则:

(1) 在所有的 3 个表中, 默认文件名是带有 .tab 扩展名的表名 (否则包含在 " " 中的文件名必须在冒号前立刻给出).

(2) 在 profit 语句中, 语法 paint<-[COL1] 告诉 AMPL, 文件 RM4profit.tab 的 COL1- 列下的条目定义了集 paint 的成员.

(3) 在 profit 语句中, 语法 unitprofit~COL2 交叉引用在 COL2 列中参数名为 unitprofit 的值.

(4) 在 rhs 语句中, [resource]IN 自动地定义集 resource 的成员, 因为名称 resource 用作表的列标题.

(5) 在 aij 语句中, aij(至少) 有两维, 因此, 不能用该语句读取相应集的成员. 相反, 这些集必须从一维表 RM4profit 和 RM4rhs 中读出. 因此, aij 语句中的 [resource, paint] 只能用于定义参数 aij 的指标. 通常, 如果模型没有指标参数,

可以声明一个表仅用于从一个文件读一个集的成员. 在这种情况下, .tab 文件中的标题行必须为 ampl.tab 1 0, 它标志文件包含一列集成员, 并且没有参数. 例如, 下面的语句声明了一个表, 用于从文件 paintSet.tab 中读集 paint 的元素.

```
table paintSet IN : [paint]IN;
```

在这种情况下, paintSet.tab 的内容是

```
ampl.tab 1 0
paint
exterior
interior
```

在一些情况下, 可以很方便地读二维参数的数据, 就像用矩阵代替两个单元素指标一样, 如前面已知的 aij. AMPL 允许这样做, 只要将表的定义改为

```
table RM4arrayAij IN:[i~resource],{j in paint}<aij[i,j]~(j)>;
```

(新表定义有点 "过编码", ~(j) 显得有点多余. 然而, 它是恰当的.) 在这种情况下, 文件 RM4arrayAij.tab 显示成

```
ampl.tab 1 2
resource exterior interior
m1            6        4
m2            1        2
demand       -1        1
market        0        1
```

注意, 标题 ampl.tab 1 2 表明, 表 RM4arrayAij 有一个关键指标 (也就是 [i~resource]) 和两个数据列 (其标题是 exterior 和 interior). 新表 RM4array-Aij 并不允许读集 resource 和 paint 的成员, 表 RM4aij 有相同的限制.(见文件 RM4.txt.)

A.5.4　输出表文件

在执行完 solve 命令之后, 表文件还可以从 AMPL 中接到输出. 其语法类似于输入文件的语法, 除了表中所声明的 IN 被 OUT 替换外. 例如, 在 Reddy Mikks 模型中, 假定我们对得到下列信息感兴趣:

(1) 变量值和它的简约费用.　　(2) 相应于约束的松弛变量和对偶价格.

这个信息需要用两张表, 因为这两项内容是不同集 paint 和 resource 的函数:

```
table varDataOUT:[paint],product,product.rc;
table conData OUT:
    [resource],limit.slack~ slack,limit.dual~Dual;
```

OUT 表声明语句应该放在约束段之后, 以保证使用在这些语句中的变量名和约束名有定义 (见文件 RM4.txt). 语法 limit.slack~slack 和 limit.dual~Dual 为在文件写入相应数据的列分配描述标题的名称 slack 和 Dual. 否则, 默认的标题名是 limit.slack 和 limit.dual.

为了将输出写到文件中, 我们需要结果命令 solve, 和紧随着的 write 语句:

```
write table varData;
write table conData;
```

将输出结果分别送到文件 varData.tab 和 conData.tab 中. 与输入情况一样, 我们可以给出 (在双引号中) 指定的文件名 (以 .tab 为扩展名) 来覆盖默认的文件名, 其文件名位置在关键词 OUT 之后和冒号之前.

可以用输出表将二维数组写到文件中. 例如, 使用下面两个定义中的任何一个均可将数组 aij 写到一个扩展名为 .tab 的文件中:

```
table AijMatrix OUT: [resource,paint], aij;
table Aijout OUT:{i in resource}->[RESOURCES],
{j in paint}<aij[i,j]~(j)>;
```

在第 1 个定义中, 文件 AijMatrix.tab 用同一行的两个指标列出 aij 的每一个元素. 在第 2 个定义中, 文件 Aijout.tab 使用数组的形式列出 aij 元素, 并以用户指定的名称 RESOURCES 作为第一 (关键) 列的标题.

习题 A.5C

1. 在文件 RM4.txt 中, 假定语句

```
read table RM4profit;
read table RM4rhs;
read table RM4aij;
```

被下面的语句替换:

```
read table RM4aij;
data;
param unitprofit:= exterior 5 interior 4;
param rhs:=m1 24 m2 6 demand 1 market 2;
```

解释为什么 AMPL 并不完全执行所作的改变. (提示: 找到答案的最佳方法是用模型作试验.)

2. 假定文件 RM4rhs.tab 所读的内容如下:

```
ampl.tab 1 1
constrName Availability
m1     24
m2     6
```

```
demand 1
market 2
```

在文件 RM4.txt 中作必要的改动, 并执行此模型.

A.5.5 电子表格形式的输入/输出表

从电子表格中读取数据和向电子表格传送数据, 这两种情况使用的语法类似于附录 A.5.4 提到的表文件的语法形式. 下面的语句说明如何从文件名为 RM5.xls 的 Excel 电子表格中读取 Reddy Mikks 模型的输入数据.

```
table profitVector IN "ODBC" "RM5.xls":paint<-[COL1], unitprofit~COL2;
table rhsVector IN "ODBC" "RM5.xls":[resource] IN,rhs;
table aijMatrix IN "ODBC" "RM5.xls":[resource,paint], aij;
```

用户命名的 `profitVector`、`rhsVector` 和 `aijMatrix` 是 Excel 表 RM5.xls 中表的名称. 这些名称定义电子表格中的区域, 相当于各自的数据表①. `"ODBC"` 是电子表格的标准数据处理界面. `read table` 则是将数据传送到模型中 (见文件 `RM5.txt`). 注意表 `profitVector` 中 COL1 和 COL2 的使用, 它们相当于电子表格中 (任意) 列的名称. 语法与输入表 (见附录 A.5.3) 是相同的. 如果希望的话, 模型的每个数据表可以储存在不同的电子表格中.

像附录 A.5.3 一样, 二维数据能够用数组格式来读, 其使用的表格定义如下:

```
table aijArray IN "ODBC" "RM5.xls":[i~resource],
{j in paint}<aij[i,j]~(j)>;
```

在这种情况下, 数组 `aij` 出现在 RM5.xls 中的区域 `aijArray` 内, 并且必须包含正确的行与列的标题. 记住, 表中所用数值型列标题必须通过使用 *Excel* 中的 `TEXT` 函数转换成字符串, 否则 AMPL 将输出一些难以辨认的错误信息.

同样的表格声明可以将输出数据传送到电子表格中. 其差别只是将 `IN` 用 `OUT` 替换, 其他的格式与表文件完全相同. 在这种情况下, `write table` 命令 (在 `solve` 命令之后) 将输出数据送到电子表格. 下面的例子演示 `OUT` 表的使用:

```
table variables OUT "ODBC" "RM5a.xls":
        [paint],product~solution,product.rc~reducedCost;
table constraints OUT "ODBC" "RM5a.xls":
        [resource],limit.slack~slack,limit.dual~dual;
```

将输出表 `variables` 和 `constraints` 传送到 Excel 文件 `RM5a.xls` 中, 在执行完 `write table` 命令后, 每个表自动出现在每一张表格的左上角.

① 为了命名一个区域, 在 Excel 表的公式栏左边的 "名称框" 中, 突出显示它并输入其名称, 然后按回车键, 或用 Excel 表中的 "插入"/"名称"/"定义".

A.6 交互式命令

AMPL 允许用户交互式地求解模型、检验/修正数据, 并将输出传送到屏幕或文件上. 下面是一些有用命令的部分列表:

```
delete 用逗号分开目标函数名和约束名;
drop 用逗号分开目标函数名和约束名;
restore 用逗号分开目标函数名和约束名;
display 用逗号分开项目列表;
print/printf 无格式/格式项目列表;
expand 用逗号分开目标函数名和约束名;
let 参数或变量 (指标或无指标) := 值;
fix 变量 (指标或无指标) := 值;
unfix 变量 (指标或无指标);
reset;
reset data;
solve;
```

这些命令在 `ampl:` 提示下交互式地输入. 一些命令, 像 `display` 和 `print`, 如果希望的话, 可以在模型中使用硬编码.

`delete` 命令完全删除所列的目标函数和/或约束, 而 `drop` 只是暂时将它们排除在模型之外. `drop` 命令可以由 `restore` 命令来废除. 新的目标函数或约束可以通过键盘输入, 将它们加入到模型中, 与我们在硬编码模型中所做的完全相同.(见例 8.2-1 对于分支定界算法的一个应用.)

我们已在 Reddy Mikks 模型中使用过 `display` 命令. 可以直接将输出传送到一个外部文件, 借助于在终止符号 (分号) 之前使用 `>文件名`. 否则, 输出的默认值是屏幕.

前面已在附录 A.5.2 中讨论过 `print/printf` 命令. 输出默认到屏幕, 也可以像 `display` 一样, 直接传送到输出文件中.

`expand` 命令提供了目标函数和约束普通形式的表达式. 例如, 在 Reddy Mikks 模型中, 命令

```
expand profit;
```

打印出的目标函数为

```
maximize profit:5*product["exterior"]+4*product["interior"];
```

用这种方法, 用户可以看到模型是否正确地得到了输入数据. 用类似的方法, 命令

```
expand limit;
```

将展开模型所有的约束. 如果你对某个特定的约束感兴趣, 则 limit 后必须加相应的指标. 例如,

```
expand limit["m1"];
```

将显示模型的第一个约束.

let 命令允许给参数或变量赋新的值 (用 := 作为赋值运算). 右端项可以是简单的数值或数学表达式. 它可以用于检验不同解的情形, 正像我们将在附录 A.7 中说明的那样.

fix 命令是在求解模型之前为变量赋特定值. 例如, 假定下列语句

```
ampl: fix product["exterior"]:=1.5;
ampl: solve;
```

在求解 Reddy Mikks 模型之前被交互式地输入. 由于这些命令, AMPL 求解问题时, 增加了限制 product["exterior"]=1.5. 由 fix 引起的变化可以利用如下的 unfix 命令撤销.

```
ampl: unfix product["exterior"];
```

当一些变量被消除 (= 0) 或者保持成常数时, fix/unfix 命令在检验模型方面非常有用. (见例 8.3-4 之后的应用到旅行商问题的 AMPL 程序.)

命令 reset 是从 AMPL 中删除当前模型的所有内容. 因此, 作为模型的重新开始, 新的 model 命令是必需的. 还有, 命令 reset data; 将删除所有的数据, 而 reset data a b c; 则只删除参数 a, b, c 的值. 再有, 在连续使用 model 命令之间, AMPL 需要 reset; 的使用. 否则会导致很难辨认的错误.

还有大量的有用的交互式命令在 AMPL 中, 但它们的详细表述已超出了本节的范围.

A.7　迭代和有条件地执行 AMPL 命令

假设在 Reddy Mikks 模型中, 我们想要研究改变指定参数下最优解的灵敏度. 例如, 在文件 RM2.txt 中, 当可利用原材料 m1 (=rhs["m1"]) 从当前值 24 吨变化到新的值 27 吨和 30 吨时, 最优解将受到什么样的影响? 在执行 RM2.txt 后, 得到 rhs["m1"]=24 的最优解, 可以进入下列交互式语句:

```
ampl: let rhs["m1"]:= 27;
ampl: solve;
ampl: display profit, product;
```

输出将显示在屏幕上 (像我们前面解释的一样, 如果希望的话, 还可以将输出送到文件中). 为获得 rhs["m1"]=30 的结果, 重复相同的语句并用 let 语句指定新的值. 然而, 这并不是做这项工作最有效的方法.

AMPL 允许建立方便的commands文件, 文件将消除重新输入命令带来的不必要的麻烦. 对于当前的例子, 命令文件 (可以指定它的文件名为 cmd.txt) 可以有如下语句:

```
for {i in 1..2}
    {
    let rhs["m1"]:=rhs["m1"]+3;
    solve;
    display profit, product;
    }
```

在模型 (此时, rhs["m1"]=24) 执行之后, 可以执行剩余的两种情况, 通过输入

```
ampl: commands cmd.txt;
```

当然, 我们也可以修正 cmd.txt, 让它包含 rhs["m1"]=24 语句. 见习题 A.7A 的第 1 题.

还可以用语句 repeat while condition {...}; 或者 repeat until condition {...}; 来代替 for{...} 语句, 其格式如下:

```
repeat while rhs["m1"]<=30
{
let rhs["m1"]:=rhs["m1"] +3;
solve;
display profit, product;
};
```

可以用另一种格式:

```
repeat until rhs["m1"]<30
{
let rhs["m1"]:=rhs["m1"]+3;
solve;
display profit,rhs["m1"], product;
};
```

注意: 只要条件的返回值为真, repeat while 就会执行循环; 而条件的返回值为假, repeat until 才执行循环.

在命令文件中另一个有用的语句是 if-then-else. 在这种情况下, if 可以跟随任意合法的条件, 而 then 和 else 只能跟随命令语句. 与 if 语句相伴的 AMPL 命令 continue; 和 break; 可以用在循环结构中, 或者跳到循环的下一个指标, 或者全部退出循环.

例 8.3-5 (图 8.14)提供了一个使用循环和条件语句来显示格式输出的很好的例子.

习题 A.7A

1. 修正 RM2.txt 使得 rhs["m1"] 假定的值在 20 吨至 35 吨之间, 其间隔为 5 吨. 所有 solve 命令必须从命令文件 cmd.txt 中执行, 其命令形式如下:

   ampl: model RM2.txt;

   ampl: commands cmd.txt;

   命令文件 cmd.txt 用下面 3 种不同形式来构造循环.

   (a) for{} (b) repeat while{}; (c) repeat until{};

A.8 用 AMPL 作灵敏度分析

在前面已经看到, 在 AMPL 线性规划模型中, 如何确定对偶价格和简约费用, 其方法是在 display 命令中使用*ConstraintName*.dual 和*VariableName*.rc. 为完成标准线性规划灵敏度分析报告, AMPL 附带提供了一种工具, 它可以确定目标函数系数的最优性区域和约束右端项 (常数) 的可行性区域. 我们将用文件 RM2.txt(见图 A.3) 来演示 AMPL 是如何生成灵敏度分析报告的.

在图 A.3 所述的模型中, 替代 solve 和 display 语句的是如下语句:

```
option solver cplex;
option cplex_options'sensitivity';
solve;
display limit.down,limit.current ,limit.up,limit.dual;
display product. down,product.current ,product.up,product.rc;
```

如果希望的话, 输出可以直接传送到文件中 (见文件 RM6.txt). 两个 option 语句必须在 solve 命令之前. 第 1 个 display 命令提供了所有约束 (在模型中称为 limit) 的可行性区域. 后缀 .down, .current 和 .up 给出了每个成员约束的右端项的下界值、当前值和上界值. 同样的方式, 第 2 个 display 命令提供了目标函数系数的最优性区域. 下列输出是很直观的.

```
profit = 21
         product.down product.current product.up product.rc
exterior    2            5               6         0
interior    3.33333      4              10         0

     limit.down limit limit.up
```

```
demand     -1.5     0       1e+20
m1         20       0.75    36
m2         4        0.5     6.66667
market     0        0       0
```

参考文献

Fourer, R., D. Gay, and B. Kernighan, *AMPL, A Modeling Language for Mathematical Programming*, 2nd ed., Brooks/Cole-Thomson, Pacific Grove, CA, 2003.

索　　引

其 他